

Self-control of traffic lights and vehicle flows in urban road networks

Stefan Lämmer¹ and Dirk Helbing^{2,3}

¹ Faculty of Transportation and Traffic Sciences 'Friedrich List', Dresden University of Technology, A.-Schubert-Strasse 23, D-01062 Dresden, Germany

² ETH Zurich, Swiss Federal Institute of Technology, UNO D 11, Universitätstrasse 41, CH-8092 Zurich, Switzerland

³ Collegium Budapest—Institute for Advanced Study, Szentháromság utca 2, H-1014 Budapest, Hungary

E-mail: traffic@stefanlaemmer.de and dhelbing@ethz.ch

Received 5 February 2008

Accepted 26 March 2008

Published 16 April 2008

Online at stacks.iop.org/JSTAT/2008/P04019

doi:10.1088/1742-5468/2008/04/P04019

Abstract. Based on fluid-dynamic and many-particle (car-following) simulations of traffic flows in (urban) networks, we study the problem of coordinating incompatible traffic flows at intersections. Inspired by the observation of self-organized oscillations of pedestrian flows at bottlenecks, we propose a self-organization approach to traffic light control. The problem can be treated as a multi-agent problem with interactions between vehicles and traffic lights. Specifically, our approach assumes a priority-based control of traffic lights by the vehicle flows themselves, taking into account short-sighted anticipation of vehicle flows and platoons. The considered local interactions lead to emergent coordination patterns such as 'green waves' and achieve an efficient, decentralized traffic light control. While the proposed self-control adapts flexibly to local flow conditions and often leads to non-cyclical switching patterns with changing service sequences of different traffic flows, an almost periodic service may evolve under certain conditions and suggests the existence of a spontaneous synchronization of traffic lights despite the varying delays due to variable vehicle queues and travel times. The self-organized traffic light control is based on an optimization and a stabilization rule, each of which performs poorly at high utilizations of the road network, while their proper combination reaches a superior performance. The result is a considerable reduction not only in the average travel times, but also of their variation. Similar control approaches could be applied to the coordination of logistic and production processes.

Keywords: flow control, traffic and crowd dynamics, traffic models, self-driven particles

Contents

1. Introduction	3
2. Network flow model	4
2.1. Traffic dynamics on road sections resulting from the continuity equation . . .	4
2.2. Kirchhoff’s law for the traffic dynamics at nodes	5
3. Anticipation of traffic flows and platoons	6
3.1. Service process and set-up times	6
3.2. Green time required to clear a queue	8
3.3. Waiting time anticipation	9
4. Conventional and self-organized traffic light control	10
4.1. The classical control approach and its limitations	10
4.2. Real-time heuristics based on a self-organized prioritization strategy	11
4.3. Optimization strategy	12
4.4. Stabilization strategy	15
4.5. Combined strategy	15
4.5.1. Service intervals.	16
4.5.2. Sufficient stability condition.	17
4.5.3. Conclusion.	19
5. Simulation of the self-organized traffic light control	19
5.1. Operation modes at an isolated intersection	20
5.1.1. Serving single vehicles at low utilizations.	20
5.1.2. Service of platoons at moderate utilizations.	21
5.1.3. Suppression of minor flows at medium utilizations.	21
5.1.4. Flow stabilization at high utilizations.	21
5.2. Performance at an isolated intersection	22
5.2.1. Constant inflows.	22
5.2.2. Variable inflows.	23
5.3. Coordination in networks	24
5.3.1. Solving the Kumar–Seidman problem.	24
5.3.2. Irregular networks.	24
6. Conclusions and outlook	26
Acknowledgments	28
Appendix. General problems of network flow coordination	28
A.1. Dynamic instabilities	28
A.2. Chaotic dynamics	29
A.3. Limited prognosis time horizon	30
References	30

J. Stat. Mech. (2008) P04019

1. Introduction

Traffic systems are a prominent example of non-equilibrium systems and have been studied extensively in the field of statistical physics [1]–[3]. Much attention was devoted to the study of self-organized phenomena in driven many-particle systems [4] such as pedestrian flows [5, 7] or traffic flows on highways [8, 9]. In order to explain phenomena like the emergence of traffic jams [10, 11] or stop-and-go waves [12]–[14], a huge variety of different traffic flow models have been proposed, e.g. follow-the-leader models [15] or fluid-dynamic traffic models in both discrete [16] and continuous [17, 18] space. More recently, a research focus was put on network traffic, which required us to extend one-dimensional traffic models in order to cope with situations, where traffic flows merge or intersect [19, 20, 7], [21]–[23]. These models can explain how jam fronts propagate backwards over network nodes [24, 25], which might eventually result in cascading breakdowns of network flows [26]–[28].

One grand challenge in this connection is the optimization of traffic lights in urban road networks [23], especially the coordination of vehicle flows and traffic lights. A typical goal is to minimize travel times, to find optimal cycle times [29, 30] and to study the corresponding spatio-temporal patterns of traffic flow [31]–[33]. It is agreed, however, that a further improvement of the traffic flow requires us to apply more flexible strategies than fixed-time controls [34]–[37]. Gershenson [38], for example, showed for a regular network with periodic boundary conditions that his control strategy synchronizes traffic lights even without explicit communication between them. Lämmer *et al* [39] proposed to represent the traffic lights by locally coupled phase oscillators, whose frequencies adapt to the minimum cycle of all nodes in the network. Further algorithms perform parameter adaptations by means of neural networks [40, 41], genetic reinforcement learning [42], fuzzy logic [43, 44] or swarm algorithms [45].

The optimization of intersecting network flows has also been studied in the domain of production [46]–[49] and control theory [50]–[53]. De Schutter and de Moor [54, 55] proposed a solution approach for finding optimal switching schedules for an isolated intersection with constant arrival rates. For networks of more than one node, Lefeber and Rooda [56] could derive a state-feedback controller from a given desired global network behavior. Besides optimality, control theorists particularly addressed the issue of the stability of decentralized control strategies [57, 58, 48]. Whereas so-called clearing policies (see appendix A.1), for example, stabilize single nodes in isolation, they might cause instabilities in networks with bidirectional flows [59]–[62]. Control strategies based on periodic switching sequences, e.g. the classical fixed-time traffic light control, have been shown to be both stable and controllable under certain conditions [63, 50].

In this paper, we propose a decentralized control algorithm, which is based on short-term traffic forecasts [64] and enables coordination among neighboring traffic lights. Rather than optimizing globally for *assumed* flow conditions that are never met exactly, we look for a heuristics that most of the time comes close to optimal operation, given the *actual* traffic situation. Assuming that it would be possible to adjust traffic regulations accordingly, we will drop the condition of periodic operation to allow for more flexible adjustment to varying traffic flows.

The fact that varying traffic flows influence the respective traffic lights ahead, which in turn influence the traffic flows, makes it *impossible* to predict the evolution of the

system over longer time horizons. This makes large-scale coordination among traffic lights difficult. It is known, however, that local nonlinear interactions can, under certain conditions, lead to system-wide spatio-temporal patterns of motion [65]. Therefore, our control concept pursues a local self-organization approach. The particular scientific challenge is that such a decentralized ‘self-control’ must be able to cope with (1) real-time optimization, (2) feedback loops due to the mutual interaction between the traffic lights via the traffic flows and (3) very limited prognosis horizons.

Our paper is organized as follows. In the next section, we introduce a fluid-dynamic model for the traffic flow in urban road networks. This model allows us to anticipate the effects of switching traffic lights (see section 3). In section 4, we explain our concept of the self-control of traffic lights. The underlying principle is inspired by the self-organization of opposite pedestrian flows, which is driven by the pressure differences between the waiting crowds. We generalize this observation in section 4.3 to define priorities of arriving traffic flows. In sections 4.4 and 4.5, the prioritization strategy is supplemented by a stabilization strategy. Simulation studies are presented in section 5 and demonstrate the superior performance of our decentralized concept of self-control.

2. Network flow model

An urban road network can be composed of links (road sections of homogeneous capacity) and nodes (intersections, merges and diverges) defining their connection. The following sections summarize a fluid-dynamic model describing the traffic dynamics on the constituents of a road network.

2.1. Traffic dynamics on road sections resulting from the continuity equation

Let us consider a homogeneous road section i with constant, i.e. time-invariant, length L_i , speed limit V_i and saturation flow Q_i^{\max} . The traffic dynamics on the road section can be characterized by the arrival rate $Q_i^{\text{arr}}(t) \leq Q_i^{\max}$ and the departure rate $Q_i^{\text{dep}}(t) \leq Q_i^{\max}$. These quantities represent the numbers of vehicles per unit time entering or leaving the road section over all its lanes.

The flow of traffic along an urban road section (in contrast to freeway sections [12]) is sufficiently well represented by Lighthill and Whitham’s fluid-dynamic traffic model [14]. It describes the spatio-temporal dynamics of congestion fronts based on the continuity equation for vehicle conservation, plus a flow density relationship known as a ‘fundamental diagram’. If we neglect net effects of overtaking and approximate the fundamental diagram by a triangular shape, this implies two distinct characteristic speeds: while perturbations of free traffic flow propagate downstream at the speed V_i , in congested traffic the downstream jam front and perturbations propagate upstream with a characteristic speed of about -15 km h^{-1} [4]. These fundamental relations also allow us to derive explicit expressions for the motion of the *upstream* jam front, where vehicles brake and enter the congested area of the road section, as well as for the related travel times [21, 18].

An integration over space results in an effective queuing-theoretical traffic model based on coupled delay-differential equations [23]. It can be summarized as follows: in free traffic, ideally, the cumulated number $N_i^{\text{exp}}(t)$ of vehicles expected to reach the downstream end

of road section i until time t is given by

$$N_i^{\text{exp}}(t) = \int_{-\infty}^t Q_i^{\text{arr}}(t' - L_i/V_i) dt', \quad (1)$$

where the time shift L_i/V_i corresponds to the travel time to pass link i in free traffic. In the case of congestion, however, the number of vehicles that have actually left the road section at its downstream end is given by the integral of the departure rate:

$$N_i^{\text{dep}}(t) = \int_{-\infty}^t Q_i^{\text{dep}}(t') dt' \leq N_i^{\text{exp}}(t). \quad (2)$$

Thus, the difference between $N_i^{\text{exp}}(t)$ and $N_i^{\text{dep}}(t)$ directly corresponds to the number of delayed vehicles, which will be referred to as the queue length $n_i(t)$. Consequently, the total waiting time $w_i(t)$ of all vehicles on road section i until time t increases at the rate

$$dw_i/dt = n_i(t) = N_i^{\text{exp}}(t) - N_i^{\text{dep}}(t). \quad (3)$$

It is important to note that, even though $n_i(t)$ does not explicitly account for the spatial location of congestion on link i , it fully captures the corresponding inflow–outflow relations, the time to resolve a queue, as well as the associated waiting times. The consistency with other and more complex traffic flow models is shown in [23].

2.2. Kirchhoff's law for the traffic dynamics at nodes

Each node in a road network connects a number of incoming road sections denoted by the index i to a number of outgoing links denoted by j . Kirchhoff's law regarding the conservation of flows at nodes requires that the flow arriving at an outgoing link j equals the sum of the fractions $\alpha_{ij}(t)$ of the departure flows $Q_i^{\text{dep}}(t)$ from the incoming links i , i.e.

$$Q_j^{\text{arr}}(t) = \sum_i \alpha_{ij}(t) Q_i^{\text{dep}}(t) \quad \text{for all } j \text{ and } t. \quad (4)$$

The turning fractions $\alpha_{ij}(t) \geq 0$ with $\sum_j \alpha_{ij} = 1$ are normalized and may be time-dependent, as route choice and travel activities can change in the course of a day [66, 4, 67, 68]. By incorporating limited arrival flows ($Q_j^{\text{arr}}(t) \leq Q_j^{\text{max}}$), it becomes obvious that a lack of arrival capacity on a downstream link limits the departure flow on the upstream links, which may eventually cause spill-back effects [69]. A discussion of concrete specifications of diverges and merges is provided in [21, 23]. For the dynamics of shock fronts propagating through such network nodes, see [70, 71].

When a traffic flow enters or crosses another one, i.e. at merging or intersection nodes, the competing traffic flows tend to obstruct each other, which often leads to an inefficient usage of intersection capacities [72, 7]. Traffic lights can serve to coordinate incompatible traffic flows and to increase the overall performance. For traffic flows served by a green light, we assume in the following that the outflow from a queue is only limited by the saturation flow Q_i^{max} . That is, throughout this paper, outflows will not be obstructed by other flows or by spill-backs from downstream road sections.

A general approach to model the switching of traffic lights is to regulate the outflow of an incoming road section i with a 'permeability' pre-factor $\gamma_i(t)$, which alternates between

0 and 1, corresponding to a red and green traffic light, respectively [23]. Three different regimes can be distinguished: (i) if the traffic light is red, the outflow is zero; (ii) when the traffic light has switched to green, the vehicle queue discharges at a more or less constant rate, the saturation flow Q_i^{\max} [73]; (iii) if the traffic light remains green after the queue has dissolved, vehicles leave link i at the same rate $Q_i^{\exp}(t) = Q_i^{\text{arr}}(t - L_i/V_i)$ at which they enter it, delayed by the free travel time L_i/V_i . Together with equation (3), one obtains an ordinary differential equation for the temporal evolution of the queue length $n_i(t)$:

$$\frac{dn_i}{dt} = \begin{cases} Q_i^{\exp}(t) & \text{if } \gamma_i(t) = 0 \\ Q_i^{\exp}(t) - Q_i^{\max} & \text{if } \gamma_i(t) = 1 \quad \text{and} \quad n_i(t) > 0 \\ 0 & \text{if } \gamma_i(t) = 1 \quad \text{and} \quad n_i(t) = 0. \end{cases} \quad (5)$$

The above model allows us to characterize the queuing process at a signalized road section as a nonlinear hybrid dynamical system [50], i.e. a system of equations containing both continuous and discrete state variables. The transition from regime (ii) to regime (iii), i.e. the transition from congested to free traffic is a result of the particular arrival flow and cannot directly be controlled by the traffic light. Thus, a complete formulation of the hybrid dynamical system requires us to anticipate the time point at which a queue will be cleared [64]. This, as well as the switching losses due to reaction times and finite accelerations, will be addressed in the following section.

3. Anticipation of traffic flows and platoons

For a flexible traffic light control to be efficient, it is essential to *anticipate the vehicle flows* as well as possible (see appendix A.3). In [64], we have proposed a framework to predict the effects of starting, continuing or terminating service processes on future waiting times. The main results are briefly summarized in the following and serve as the basis for deriving optimal switching rules in section 4.3.

Note, however, that there are fundamental limits to the prediction of traffic flows (see appendix). Already very small networks with very simple switching rules can produce a complex and potentially chaotic traffic dynamics (see appendix A.2). Moreover, coordination problems between traffic flows and their service may cause an inefficient usage of intersection capacities and, thereby, spill-back effects and related dynamic instabilities (see figures 1 and 2 and appendix A.1). These can sometimes be quite unexpected and imply that plausible optimization attempts may fail due to nonlinear feedback effects. Details are discussed in the appendix.

3.1. Service process and set-up times

The safe operation of traffic lights requires that, before switching to green for the traffic flow of i , all other incompatible traffic flows have been stopped and all corresponding vehicles have already left the conflict area. This will be considered in our model by introducing set-up times: if some traffic flow i is selected for service, its traffic light does not switch to green before the corresponding set-up (or intergreen) time τ_i^0 has elapsed [74]. The value of τ_i^0 has to be chosen according to safety considerations and usually lies in the range between 3 and 8 s. Please note that τ_i^0 also includes the amber time period, which takes into account reaction delays and delays by finite acceleration. Therefore, the set-up

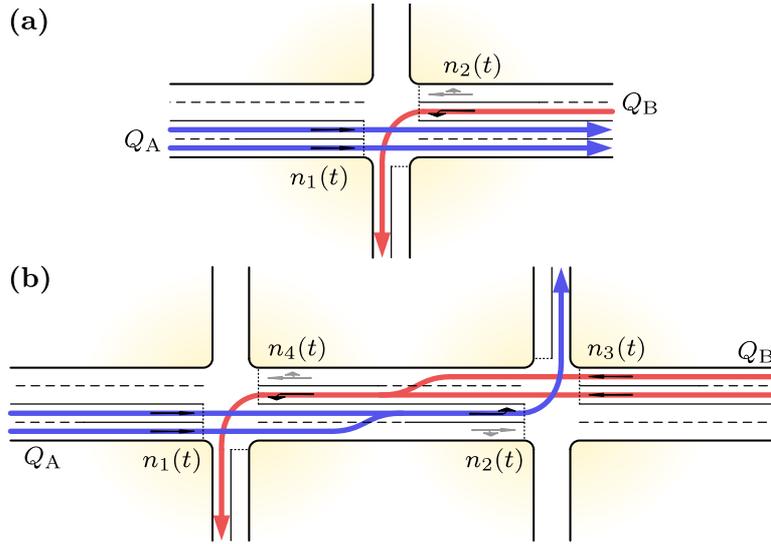


Figure 1. (a) Isolated intersection with two incompatible traffic streams A and B. In this case, a suitable clearing policy is both optimal and stable (see appendix A.1). (b) Combination of two intersections of the kind displayed in (a), forming a non-acyclic road network (see appendix A.1). It is interesting that, even when each of the intersections behaves stably in isolation, the road network might behave dynamically unstable under identical inflow conditions (see figure 2 and [59]).

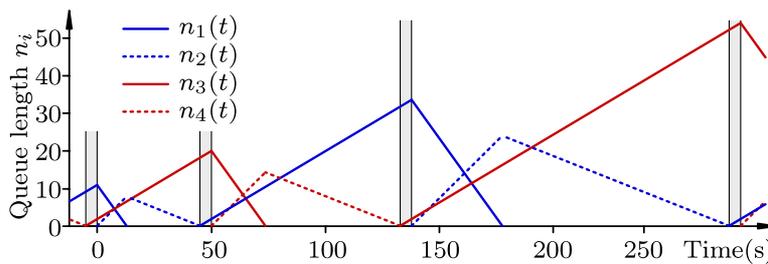


Figure 2. Time-dependent queue lengths for the non-acyclic road network shown in figure 1(b), assuming a clearing policy that behaves optimally at the isolated intersection illustrated in figure 1(a). The queue lengths diverge due to dynamic instability. (For an explanation of the clearing policy see appendix A.1.) The reason for this instability lies in the inefficient usage of service capacities during the time periods from 20 to 45 s, from 70 to 130 s, and so on. During this time, the traffic lights extend the green time for streets 1 and 3 where the vehicle queues have already been cleared, while the other streets are ‘being starved of input’, using the words of Kumar and Seidman [59].

time τ_i^0 reflects all time losses associated with the start of service for vehicles on link i . As depicted in figure 3(c), a service process can be divided into three successive states: the set-up, the clearing of the queue and the green time extension. The traffic light is green only in the latter two states.

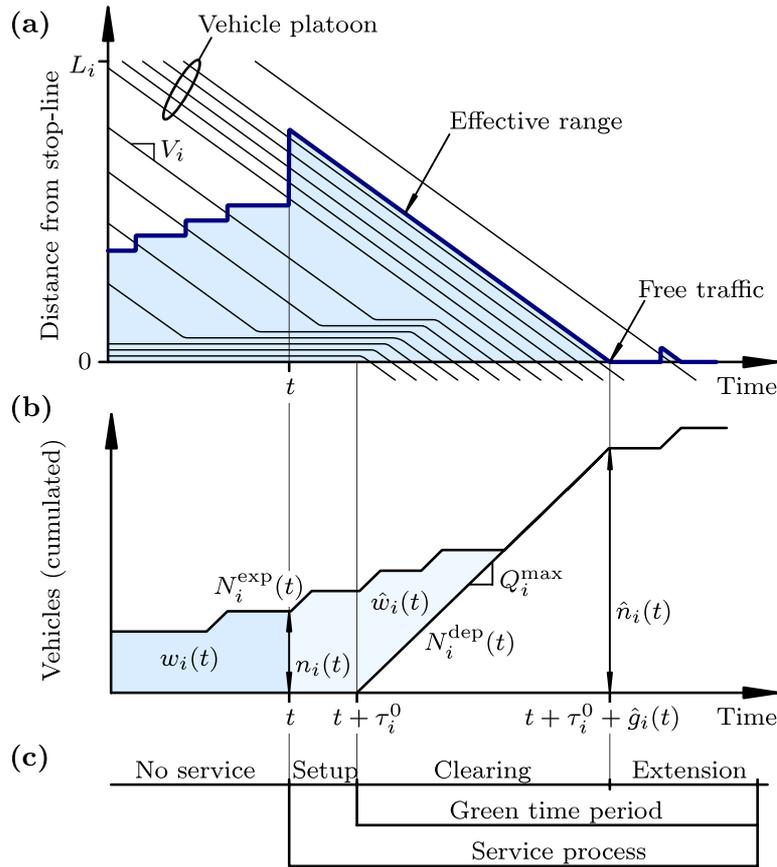


Figure 3. (a) Trajectories and (b) cumulated number of vehicles on a road section i , and (c) different states of the service process. The service process starts early enough to serve a platoon of five vehicles in a green-wave manner, i.e. without stopping the vehicles. The precise timing results from a short-term anticipation [64] based on the time series $N_i^{\text{exp}}(t)$ and $N_i^{\text{dep}}(t)$ (i.e. the cumulated number of vehicles that could have reached the stop-line in free traffic as compared to the number that actually have left the road section, see equations (1) and (2)). Whereas the current waiting time $w_i(t)$ grows with the number of vehicles $n_i(t)$ being delayed (see equation (3)), the expected future waiting time $\hat{w}_i(t)$ grows with the expected number of vehicles $\hat{n}_i(t)$ to be served in the subsequent ‘clearing’ state (see equations (8) and (9)). The value of $\hat{n}_i(t)$ as well as the required green time $\hat{g}_i(t)$ for clearing the queue are determined by equations (6) and (7). A platoon is served in a green-wave manner, if the start of the service process is initiated by the platoon-related jump in $\hat{n}_i(t)$ or, what is more illustrative, by the sudden increase of the effective range (see text).

3.2. Green time required to clear a queue

For the flexible control of traffic lights it is of fundamental importance to anticipate the amount of green time $\hat{g}_i(t)$ required for clearing the queue in road section i , given the service starts or is being continued at the current time point t . Obviously, $\hat{g}_i(t)$ does not only depend on the current queue length $n_i(t)$, but also on the number of vehicles joining

the queue during the remaining set-up time $\tau_i(t)$ and while the queue is being cleared. The queue of delayed vehicles has fully dissolved at the time point $t + \tau_i(t) + \hat{g}_i(t)$, which is defined by the requirement that the number of vehicles having left the road section by that time is equal to the number of vehicles that have reached the stop-line. This corresponds to the left- and right-hand side, respectively, of the following equation:

$$N_i^{\text{dep}}(t) + \hat{g}_i(t) Q_i^{\text{max}} = N_i^{\text{exp}}(t + \tau_i(t) + \hat{g}_i(t)). \quad (6)$$

The value of $\hat{g}_i(t)$ shall be the largest possible solution of equation (6), which can be easily obtained with standard bisection methods [75]. The second term in equation (6) represents the number of vehicles that are expected to leave the road section at the saturation flow rate Q_i^{max} , and shall be denoted by $\hat{n}_i(t)$, i.e.

$$\hat{n}_i(t) = \hat{g}_i(t) Q_i^{\text{max}}. \quad (7)$$

A detailed derivation and discussion of the dynamics of $\hat{g}_i(t)$ and $\hat{n}_i(t)$ in different dynamical regimes is provided in [64]. $\hat{n}_i(t)$ captures all those vehicles

- already waiting in the queue,
- joining the queue during set-up or clearing and
- arriving as a platoon immediately after the queue is cleared.

It particularly considers jumps to a higher value, when a platoon could be served in a green-wave manner, i.e. without stopping. The magnitude of the jump is equal to the size of the platoon. Before the platoon arrives at the stop-line, the formula reserves exactly as much time as needed to perform the set-up and to clear the queue of waiting vehicles. Thus, the above anticipation model provides us with a mechanism that establishes green waves. In order to visualize the underlying principle, figure 3(a) plots the so-called effective anticipation range, which includes the $\hat{n}_i(t)$ vehicles.

Note that, when the effective range extends $(\tau_i(t) + \hat{g}_i(t))V_i$ meters from the stop-line, all vehicles within that range will reach the stop-line *before* the queue is being cleared at time point $t + \tau_i(t) + \hat{g}_i(t)$. These vehicles will thus be served within the ‘clearing’ state of a *subsequent* service process.

3.3. Waiting time anticipation

Obviously, we would like to be able to decide whether to continue a service process or start another one is more profitable in terms of saving waiting time. Therefore, the above anticipation concept shall now be used to forecast the total waiting time $\hat{w}_i(t)$ of all vehicles on road section i up to the end of the subsequent ‘clearing’ state (see figure 3(b)). According to [64], we have

$$\frac{d\hat{w}_i}{dt} = \begin{cases} \hat{n}_i(t) & \text{if } i \text{ is not served} \\ 0 & \text{during the entire service process.} \end{cases} \quad (8)$$

That is, any delay dt in the start of service will cause an additional delay dt for each of the expected vehicles. Interestingly, $\hat{w}_i(t)$ does not change anymore during the service process, because the corresponding value has already been anticipated before. However, it will change again as soon as the service process is terminated. At the same point in time, the anticipated waiting time $\hat{w}_i(t)$ will also increase by the additional amount $\Delta\hat{w}_i(t)$

due to the fact that the next green time cannot start before performing a new set-up, which takes a time period τ_i^0 . This additional set-up waiting time is given by

$$\Delta \hat{w}_i(t) = Q_i^{\max} \int_{\tau_i(t)}^{\tau_i^0} \hat{g}_i(t, \tau') d\tau', \quad (9)$$

where $\hat{g}_i(t, \tau)$ corresponds to the solution of equation (6), given a remaining set-up time of τ' . The above equations (8) and (9) allow one to anticipate the costs of delaying or terminating a service process in terms of expected future waiting times. To underline the particular importance of this result, we would like to point out the direct relation between $n_i(t)$ and $\hat{n}_i(t)$: while $n_i(t)$ is the growth rate of the *current* waiting time $w_i(t)$ according to equation (3), $\hat{n}_i(t)$ is the growth rate of the *expected* future waiting time $\hat{w}_i(t)$ for a traffic flow i that is not being served. This fundamental similarity allows us to easily transfer conventional control schemes, which have originally been developed to operate on $n_i(t)$, to the variables of our anticipation model.

4. Conventional and self-organized traffic light control

4.1. The classical control approach and its limitations

The optimal control of switched network flows is known to be an NP-hard problem [76], which means that the time required to find an optimal solution grows faster than polynomially with the network size (number of nodes). This NP-hardness has two major implications: first, traffic light controls for road networks are usually optimized off-line for certain standard situations (such as the morning or afternoon rush hours, sports events, evening traffic, weekends, etc), and applied under the corresponding traffic conditions. Second, today's control approaches are predominantly centralized and based on the application of pre-calculated periodic schedules, some parameters of which may be adaptively adjusted (for a discussion of the related traffic engineering literature see [77, 78]). That is, coordination is reached by applying a common cycle time to all intersections or multiples of a basic frequency [73]. This frequency is normally set by the most serious bottleneck. For capacity reasons (to minimize inefficiencies due to switching times), the frequency is reduced at high traffic volumes, but it is limited by a maximum admissible cycle time. Apart from the cycle time, the order and relative duration of green phases (the 'split') and the time shifts between neighboring traffic lights ('offsets') are optimized for assumed boundary conditions (inflow and outflow). The resulting program usually serves each traffic flow once during the cycle time and it is repeated periodically. So-called 'green waves' are implemented by suitable adjustment of green phases and time shifts. They usually prioritize a unidirectional main flow (e.g. in- or out-bound rush-hour traffic in 'arterials') [79].

Some obvious disadvantages of this classical control approach are:

- (i) In order to cope with variations of the inflow, green times are often longer than needed to serve the average number of arriving vehicles (otherwise excessive waiting times may occur due to multiple stops in front of the same red light). This causes unnecessarily long waiting times for incompatible flow directions.
- (ii) At intersections with small utilization, the cycle time is typically much longer than required (or the cycle is uncoordinated with the intersection constituting the major

bottleneck). Moreover, traffic lights tend to cause avoidable delays during times of light traffic (e.g. at night).

- (iii) A coordination through ‘green waves’ is applicable to one traffic corridor and flow direction only, while they tend to obstruct opposite, crossing and merging flows.
- (iv) Due to the considerable variation of traffic flows and turning fractions from one minute to another, the traffic light schedule is optimized for an *average* situation which is never met exactly, while it is not optimal for the *actual* traffic situation.

4.2. Real-time heuristics based on a self-organized prioritization strategy

To overcome the previously mentioned disadvantages, we propose to perform a heuristic on-line optimization that flexibly adapts to the *actual* traffic situation at *each* time and place. If this heuristic reaches, on average, say 95% of the performance of the theoretically optimal solution, it is expected to be superior to the pre-determined 100% best solution for an *average* traffic situation that never occurs exactly. Moreover, finding the one 100% best traffic light control for a given time-dependent situation is numerically so demanding that it requires off-line optimization, while solutions reaching, say, 95% of the optimal performance can be determined in real time. As there are typically *several* alternative solutions of high, but not optimal, performance it is also possible to select a solution that is particularly well adjusted to the *local* traffic conditions.

In the following, we will specify a heuristics for a decentralized, real-time traffic light control. In order to reach a superior performance as compared to a simple, cyclical fixed-time control (see section 5), our self-organized prioritization approach combines an optimizing strategy (see section 4.3) with a stabilizing one (see section 4.4). Our concept goes beyond adaptive traffic light controls based on simple feedback between the traffic situation and the traffic control. It has the characteristic features of self-organized systems:

- (i) openness and autonomy,
- (ii) the lack of a central plan or hierarchical control,
- (iii) complexity, i.e. non-stationary and typically even non-periodic dynamics due to nonlinear interactions between varying interaction partners and
- (iv) no distinction between controlling and controlled elements (here: traffic lights control vehicles, and these influence the traffic lights).

The phenomena observed (such as the emergence of green waves, and others) spontaneously result from the interactions in the system, and discontinuous transitions are expected to take place among the various occurring operation regimes (see section 5.1). This is fundamentally different from the classical ‘adaptive control’ concept, where a clear distinction is made between controlled and control elements, and where control parameters are eventually adapted to changing (e.g. exponentially averaged) traffic conditions.

Our control concept is inspired by the observation that pedestrian counter-flows at bottlenecks show a self-organized oscillation of their passing direction (see figure 4), as if the pedestrians were controlled by traffic lights [5, 80, 7]. In pedestrian flows, the self-organized oscillations result from pressure differences between the waiting crowds on both sides of the bottleneck. Pressure builds up on the side where more and more pedestrians have to wait, while it is reduced on the side where pedestrians manage to pass the

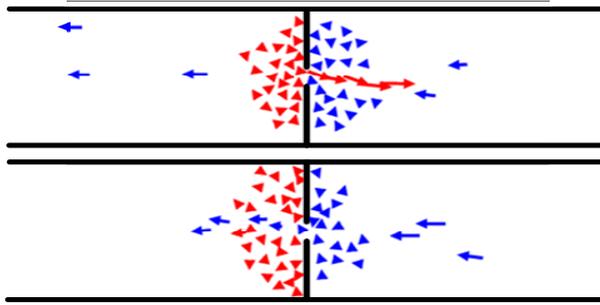


Figure 4. Pedestrians flows at a narrow bottleneck behave almost as if they were controlled by traffic lights (after Helbing and Molnár [5]).

bottleneck. The passing direction changes when the pressure on one side exceeds the pressure on the other side by a sufficient amount.

Intersections may also be viewed as bottlenecks, but with more than two flows competing for the available service capacity. Therefore, we had the idea to transfer the above-described self-organizing principle to urban vehicular traffic, although the analogy is certainly limited. For example, we find an oscillatory queue formation process behind the intersection bottleneck, but in contrast to pushy pedestrians with frictional interactions [6], we do not observe arching and clogging effects.

4.3. Optimization strategy

In our network traffic flow model, we define ‘pressures’ by dynamic priority indices $\pi_i(t)$ such that the traffic lights of an intersection give a green light to the traffic flow i with highest priority. For the mathematical formulation of the dynamic prioritization rule, let us store the argument i in a decision variable $\sigma(t)$ as follows:

$$\sigma(t) = \arg \max_i \pi_i(t). \quad (10)$$

Priority-based scheduling has been studied in the context of queuing theory [49], [81]–[87]. It has been stated that ‘there are no undiscovered priority index sequencing rules for minimizing total delay costs’ [82]. However, the considered prioritization strategies were restricted to functions of the *current* queue length, i.e. to the number of vehicles that have *already* been stopped [81, 88]. In contrast, our anticipation model (see section 3) allows one to predict *future* arrivals and to generalize these strategies to serving platoons without any previous stops, i.e. in a ‘green-wave’ manner. For simplicity, we will assume in the following that route choice is non-adaptive (i.e. the turning fractions $\alpha_{ij}(t)$ are known) and also that all traffic flows at the intersections are conflicting (i.e. only one traffic flow can be served at a time).

Our goal is to derive a formula for the priority index π_i such that switching rule (10) minimizes the total waiting time. However, the optimization horizon is limited to those vehicles, whose future waiting time directly depends on the *current* state of the traffic lights, i.e. the expected \hat{n}_i vehicles captured within the effective range (see figure 3(a)). Later arriving vehicles are neglected as long as they are beyond the anticipation horizon, but they are taken into account by the dynamic re-optimization early enough to serve them by a green wave if this is possible.

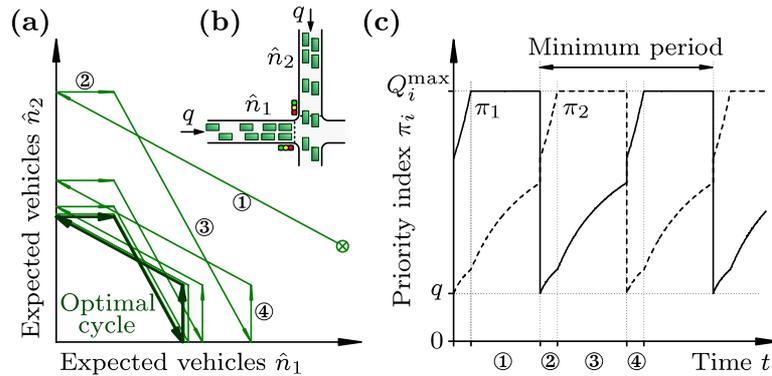


Figure 5. (a) Convergence of the trajectories (\hat{n}_1, \hat{n}_2) to the optimal limit cycle at an intersection with two identical traffic flows with constant inflow rate q , see (b). (c) Periodic time series of the priority indices π_1 and π_2 associated with the optimal cycle. (① means clearing street 1, ② set-up for street 2, ③ clearing street 2 and ④ set-up for street 1).

In case of no further arrivals, Rothkopf and Smith [82] showed that the optimal order of serving traffic flows is unique and can be determined by comparing priorities among *pairs* of competing traffic flows. This allows us to derive the optimal specification of the priority index π_i by studying an intersection of only *two* competing traffic flows 1 and 2, as depicted in figure 5(b). For the current time point t , we assume the remaining set-up times τ_1 and τ_2 , the anticipated number of vehicles \hat{n}_1 and \hat{n}_2 , and the required green times \hat{g}_1 and \hat{g}_2 to be given. We assume that, initially, traffic flow 1 is being selected for service, i.e. $\sigma = 1$. In this scenario, the controller has two options:

- (i) to finish serving flow 1 before switching to flow 2 or
- (ii) to switch to flow 2 immediately, at the cost of an extra set-up for switching back to flow 1 later on.

The optimal control decision is derived by calculating the total increase in the anticipated waiting time for each option. Following the first option requires continuing to serve flow 1 for $\tau_1 + \hat{g}_1$ seconds. According to equation (8), the anticipated waiting time of traffic flow 2 grows at the rate \hat{n}_2 , while it remains constant for traffic flow 1 under service. Since it also does not change after queue 1 has been cleared and while flow 2 is being served, the total increase of the anticipated waiting time associated with the first option would be

$$(\tau_1 + \hat{g}_1) \hat{n}_2. \quad (11)$$

When selecting the second option, according to equation (9) the termination of the service of traffic flow 1 causes the anticipated waiting time to increase by the amount $\Delta\hat{w}_1$, which reflects the extra waiting time associated with the set-up for switching back later. While serving traffic flow 2 for $\tau_2 + \hat{g}_2$ seconds, the anticipated waiting time grows further at the rate \hat{n}_1 . Altogether, its total increase would be

$$\Delta\hat{w}_1 + (\tau_2 + \hat{g}_2) \hat{n}_1. \quad (12)$$

Thus, it is optimal to continue serving traffic flow 1 as compared to switching to flow 2 if

$$(\tau_1 + \hat{g}_1) \hat{n}_2 < \Delta \hat{w}_1 + (\tau_2 + \hat{g}_2) \hat{n}_1. \quad (13)$$

The above optimality criterion allows us to define priority indices π_1 and π_2 by separating the corresponding variables. For this, we rewrite equation (13) in the following way:

$$\pi_1 := \frac{\hat{n}_1}{\tau_1 + \hat{g}_1} > \frac{\hat{n}_2}{\Delta \hat{w}_1 / \hat{n}_1 + \tau_2 + \hat{g}_2} =: \pi_2. \quad (14)$$

Each side of this inequality defines a priority index π_i . With this definition, the priority π_1 for traffic flow 1 is a function of its own variables only. Interestingly, π_2 has the same dependence on its own variables, but it additionally depends on the term $\Delta \hat{w}_1 / \hat{n}_1$. Before we can derive a general formula for the priority π_i of *any* traffic flow i , we must first clarify the role of this extra term. In general, the expression $\Delta \hat{w}_\sigma / \hat{n}_\sigma$ reflects the penalty for terminating the current service process, where σ stands for the traffic flow being served. As follows from equation (9), the value of $\Delta \hat{w}_\sigma / \hat{n}_\sigma$ ranges from 0 to τ_σ^0 and thus represents the additional waiting time $\Delta \hat{w}_\sigma$ due to the extra set-up for switching back, averaged over all corresponding vehicles \hat{n}_σ . Since the penalty for switching from σ to i applies only to those traffic flows $i \neq \sigma$ not being served, we can introduce the general penalty term $\tau_{i,\sigma}^{\text{pen}}$ as follows:

$$\tau_{i,\sigma}^{\text{pen}} = \begin{cases} \Delta \hat{w}_\sigma / \hat{n}_\sigma & \text{if } i \neq \sigma \\ 0 & \text{if } i = \sigma. \end{cases} \quad (15)$$

With this notation we can introduce the general definition of the priority index π_i as

$$\pi_i = \frac{\hat{n}_i}{\tau_{i,\sigma}^{\text{pen}} + \tau_i + \hat{g}_i}. \quad (16)$$

This is fully compatible with the optimality criterion (14). To interpret the result, the priority index π_i relates to the anticipated average service rate, i.e. the anticipated number \hat{n}_i of vehicles expected to be served during the time period $\tau_i + \hat{g}_i$. In contrast to conventional priority specifications derived from the so-called μc rule [81, 83, 88, 89], specification (16) is novel in two fundamental aspects: first, its dependence on the predicted variables \hat{n}_i and \hat{g}_i allows one to anticipate future arrivals (see section 3). Second, it takes into account both first- and second-order switching losses, i.e. the set-up times for switching to another traffic flow as well as for switching back, represented by τ_i and $\tau_{i,\sigma}^{\text{pen}}$, respectively.

Instead of clearing existing queues in the most efficient way, our anticipative prioritization strategy aims at minimizing waiting times. This prevents queues to form and causes green waves to emerge automatically, whenever this saves overall waiting time at the intersection. The underlying mechanism relates to the fact that the values of \hat{n}_i and \hat{g}_i jump to a higher value as soon as the first vehicle of a platoon enters the dynamic anticipation horizon (see section 3).

Whether a platoon is being served by a green wave or not finally depends, of course, on the overall traffic situation at the local intersection. While our previous considerations applied to vehicle queues of given length, the same prioritization rule shows a fast exponential convergence to the optimal traffic light cycle also for continuous

inflows (see figure 5). However, a local optimization of each single intersection must not necessarily imply global optimality for the entire network [56], [90]–[92], as dynamic instabilities cannot be excluded (see appendix A.1). Thus, our self-organized traffic light control must be extended by a stabilization strategy.

4.4. Stabilization strategy

We call a traffic light control ‘stable’ if the queue lengths will always stay finite [57]. Of course, stability requires that the traffic demand does not exceed the intersection capacities. Nevertheless, the short-sightedness of locally optimizing strategies could lead to an inefficient use of capacity, e.g. because of too frequent switching or too long green time extensions. This problem can be illustrated even by analytical examples, see [59, 88, 48]. For a discussion see section 3 and appendix A.1. As a consequence, even when the traffic demand is far from being critical, there is a risk that vehicle queues grow longer and longer and eventually block traffic flows at upstream intersections [26].

In order to stabilize a switched flow network, one may implement local supervisory mechanisms [59]. The function of such mechanisms is to observe the current traffic condition and to assign sufficiently long green times before queues become too long. Maintaining stability is more of a resource allocation (green time assignment) rather than a scheduling problem.

Our proposal is to complement the prioritization rule (16) by the following stabilization rule: we define an ordered priority set Ω containing the arguments i of all *those* traffic flows that have been selected by the supervisory mechanism and, thus, need to be served soon in order to maintain stability. Furthermore, the argument i of a crowded link i *joins* the set Ω as soon as more than some critical number n_i^{crit} of vehicles is waiting to be served. It is *removed* from the set after the queue was cleared, i.e. $n_i = 0$, or after a maximum allowed green time g_i^{max} was reached. Elements included in the set Ω are served on a first-come-first-served basis. As long as Ω is not empty, the control strategy is to always serve the traffic flow corresponding to the first element (head) of Ω . If Ω is empty, the traffic lights follow the prioritization rule (10).

4.5. Combined strategy

Our new control strategy can be summarized as follows:

$$\sigma = \begin{cases} \text{head } \Omega & \text{if } \Omega \neq \emptyset \\ \arg \max_i \pi_i & \text{otherwise.} \end{cases} \quad (17)$$

It is, therefore, a combination of two complementary control regimes. Whereas the optimizing regime (while $\Omega = \emptyset$) aims for minimizing waiting times by serving the incoming traffic as quickly as possible, the stabilizing regime (while $\Omega \neq \emptyset$) intervenes only if the optimizing regime fails to keep the queue lengths below a certain threshold n_i^{crit} . This means that, as long as the optimizing regime itself exhibits the desired behavior, i.e. as long as it is stable, the stabilizing regime will *never* intervene. If it needs to be activated for particular traffic flows i with $\hat{n}_i > n_i^{\text{crit}}$, however, the control is handed back to the optimizing regime as soon as the critical queues have been cleared.

Originally, such stabilizing supervisory mechanisms have been proposed for the control of production and communication systems, e.g. in [59, 88], [93]–[95]. As such rules would,

however, not explicitly pay attention to the duration of red traffic lights, they would not be suited for the application to urban road networks: too long red times would increase the risk of red-light violations and therefore also the risk of traffic accidents [96]–[98]. Thus, it is essential to have a good model for the service intervals.

4.5.1. Service intervals. In the following, we will specify the critical thresholds n_i^{crit} and the maximum green times g_i^{max} such that the stabilization rule alone ($\sigma = \text{head } \Omega$) fulfills the following two safety requirements: each traffic flow shall be served

- (S1) once, on average, within a desired service interval $T > 0$ and
- (S2) at least once within a maximum service interval $T^{\text{max}} \geq T$.

These two parameters, T and T^{max} , are the only two adjustable parameters of our control algorithm.

As service interval z_i , we define the time interval between two successive service processes for the same traffic flow i . Accordingly, the service interval z_i is the sum

$$z_i = r_i + \tau_i^0 + \hat{g}_i \tag{18}$$

of the preceding red time of r_i , the set-up time τ_i^0 and the green time \hat{g}_i anticipated before the start of the service process. Thus, we can anticipate the service interval z_i before the corresponding service process starts. This allows us to replace the critical threshold n_i^{crit} by a function $n_i^{\text{crit}}(z_i)$ of the anticipated service interval z_i .

Let us now study the statistical distribution of the service interval z_i for a traffic flow i with random arrivals. Under the assumption that $n_i^{\text{crit}}(z_i)$ is non-increasing and the traffic flow is being served as soon as $\hat{n}_i \geq n_i^{\text{crit}}(z_i)$, we can make the following general statement: the probability $P(Z \leq z_i)$ that the service interval Z is shorter than z_i is equal to the probability that more than $n_i^{\text{crit}}(z_i)$ vehicles arrive within a time interval z_i . The probability distribution $P(Z \leq z_i)$ can be derived from a given function $n_i^{\text{crit}}(z_i)$ and a given stochastic model of the arrival process, for example using the framework proposed in [99, 100]. Figure 6 illustrates the distribution for two different threshold functions $n_i^{\text{crit}}(z_i)$.

From the above observations, we can now derive an appropriate specification of $n_i^{\text{crit}}(z_i)$. Most importantly, safety requirement (S1) can be fulfilled independently of the particular arrival process. Following the above arguments, this means that $P(Z \leq T^{\text{max}}) = 1$ can be enforced by requiring

$$n_i^{\text{crit}}(z_i) \leq 0 \quad \text{for } z \geq T^{\text{max}}. \tag{19}$$

Thus, no matter how few vehicles actually arrived, the corresponding traffic flow will be served once within T^{max} . One possible specification is

$$n_i^{\text{crit}}(z_i) = \bar{Q}_i T \frac{T^{\text{max}} - z_i}{T^{\text{max}} - T}, \tag{20}$$

where \bar{Q}_i denotes the average arrival rate. This specification satisfies condition (19), but also fulfills the safety requirement (S2). Within the desired service interval T , there will, on average, arrive a number of $\bar{Q}_i T$ vehicles. This number, however, is equal to the critical threshold $n_i^{\text{crit}}(z_i)$ for an anticipated service interval of $z_i = T$. Thus, a service process is started immediately when there are as many vehicles to serve, as there arrive

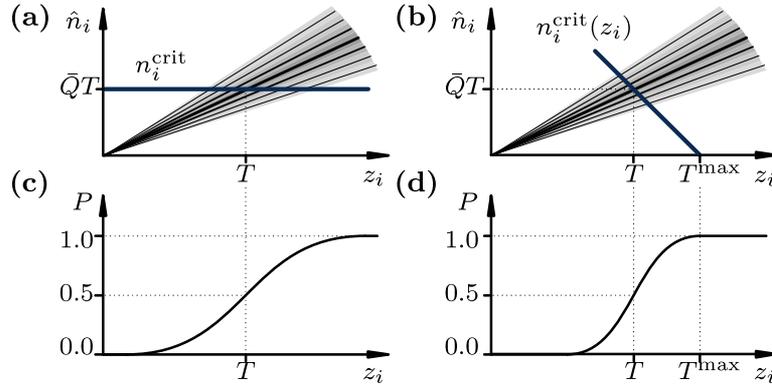


Figure 6. Top: anticipated number of vehicles \hat{n}_i to be served within a service interval z_i , following from the stochastic arrivals of the vehicles (fan of diagonal lines). If the service process is started as soon as \hat{n}_i exceeds the threshold function $n_i^{\text{crit}}(z_i)$ (thick line), the corresponding service interval is z_i . $n_i^{\text{crit}}(z_i)$ was specified as in equation (20) and plotted for $T^{\text{max}} \rightarrow \infty$ (left) and $T^{\text{max}} < \infty$ (right). Bottom: probability distribution $P(Z \leq z_i)$. Because the probability for $z_i < T$ is 50%, and for $z_i < T^{\text{max}}$ it is 100%, our controller fulfills both safety requirements, (S1) and (S2).

on average within the desired service time period T . Figure 6 plots the distribution of service intervals z_i for different parameters of the threshold function $n_i^{\text{crit}}(z_i)$ according to specification (20). Altogether, the probability of having $z_i < T$ is 50% and the probability for $z_i < T^{\text{max}}$ is 100%.

Let us briefly discuss two limiting cases: (i) if $T^{\text{max}} \rightarrow \infty$, the threshold function $n_i^{\text{crit}}(z_i) = \bar{Q}_i T$ becomes a horizontal line as depicted in figure 6(a). This parameter choice corresponds to a fully vehicle-responsive operation, where one does not care about the duration of the actual service interval. (ii) If $T^{\text{max}} \rightarrow T$, the threshold function becomes a vertical line at $z_i = T$. This case, in contrast, corresponds to a pure fixed-time operation with cycle time T , where the actual traffic situation is completely ignored. In between these two limiting cases, i.e. for $T < T^{\text{max}} < \infty$, the switching behavior is both time-dependent and vehicle-responsive.

4.5.2. Sufficient stability condition. To make our control concept complete, the last step is to specify the maximum allowed green time g_i^{max} for traffic flow i in the stabilization strategy. Once g_i^{max} is exceeded, the element i is removed from Ω , even if its queue has not been fully cleared in this time. Obviously, g_i^{max} must be chosen large enough in order to maintain stability [59]. In particular, serving an average number of $\bar{Q}_i T$ vehicles requires us to provide a green time of at least $T\bar{Q}_i/Q_i^{\text{max}}$ seconds. On the other hand, serving all traffic flows one after the other for $\tau_i^0 + g_i^{\text{max}}$ seconds each should not take more than T seconds in total. Therefore, g_i^{max} must meet the constraints

$$g_i^{\text{max}} \geq T \bar{Q}_i / Q_i^{\text{max}} \quad \text{for all } i \quad (21)$$

and

$$\sum_i (\tau_i^0 + g_i^{\text{max}}) \leq T. \quad (22)$$

In order to obtain a sufficient condition for the existence of stable solutions, one can insert g_i^{\max} from equation (21) into equation (22), which leads to

$$\sum_i \tau_i^0 \leq \left(1 - \sum_i \bar{Q}_i / Q_i^{\max}\right) T. \quad (23)$$

That is, the sum of set-up times must be smaller than the fraction of the service period T not needed to serve arriving vehicles. This condition is consistent with the condition of Savkin [63, 101] for a general switched server queuing system to be controllable. Condition (23) also indicates that there is a lower threshold for the desired service time period T :

$$T \geq \frac{\sum_i \tau_i^0}{1 - \sum_i \bar{Q}_i / Q_i^{\max}}. \quad (24)$$

Interestingly, the same threshold has been shown to be the shortest possible cycle for a stable *periodic* switching sequence [73, 50, 102, 39]. Therefore, we can conclude that our self-organized, non-periodic traffic control defined by equation (17) is stable whenever there exists a stable *fixed-time* control with cycle time T .

For a given desired service time period T satisfying the stability condition (24), the corresponding g_i^{\max} values can be obtained by solving an optimization problem. To minimize the average waiting times over an interval T , one maximizes the overall throughput $\sum_i Q_i^{\max} g_i^{\max}$ as proposed in [103, 104, 55]. In order to solve this optimization problem, however, it is necessary to know how much green time must be reserved for all other traffic flows [35]. The determination of the exact optimum would require us to predict future arrivals over a prognosis horizon of about T (i.e. normally much longer than one minute). Because this is usually not possible (see appendix A.3), we suggest to determine a nearly optimal solution instead. Setting

$$g_i^{\max} = \frac{\bar{Q}_i}{Q_i^{\max}} T + \frac{Q_i^{\max}}{\sum_{i'} Q_{i'}^{\max}} T^{\text{res}} \quad (25)$$

(see the circle in figure 7) satisfies both constraints (21) and (22). The first term on the right-hand side of equation (25) represents the minimum required green time $T\bar{Q}_i/Q_i^{\max}$ according to equation (21). The second term adds a fraction of the ‘residual time’ T^{res} proportional to the corresponding saturation flow Q_i^{\max} . Herein, the residual time is defined as

$$T^{\text{res}} = T \left(1 - \sum_i \bar{Q}_i / Q_i^{\max}\right) - \sum_i \tau_i^0, \quad (26)$$

i.e. as the part of the service interval T that is not necessarily needed for service processes. (In other words, stability would still be guaranteed even if the traffic lights would not serve any traffic flow for T^{res} seconds within the service interval T . Thus, T^{res} relates to the free intersection capacity, which is here being used to provide maximum possible green times if they are needed by the stabilization strategy.)

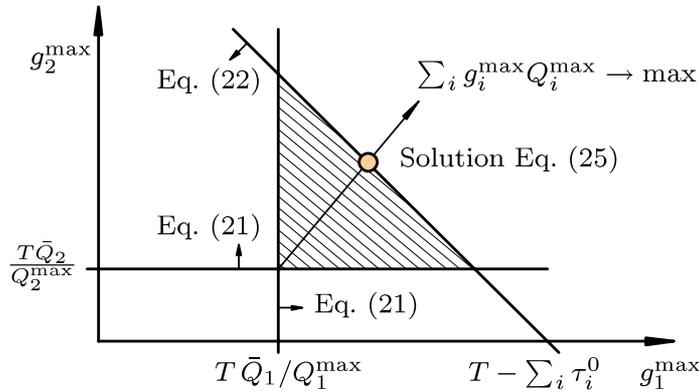


Figure 7. Stable solutions for the maximum green times g_i^{\max} lie within the simplex (shaded area) constrained by equations (21) and (22). The optimal values for g_i^{\max} can be obtained from maximizing the throughput $\sum_i g_i^{\max} Q_i^{\max}$. An easily computable explicit solution (circle) is given by equation (25).

4.5.3. Conclusion. Our decentralized traffic light control strategy given by equation (17) should stabilize traffic flows in a road network as long as the traffic demands \bar{Q}_i and the desired service interval T satisfy the sufficient stability condition (24). Interestingly, this condition is satisfied whenever there exists a stable fixed-time control with cycle time T . Furthermore, both safety requirements (S1) and (S2) are fulfilled under all circumstances, i.e. even for over-saturated traffic conditions, where the eventual growth of vehicle queues is unavoidable. In this case, the stabilization strategy serves the ingoing traffic flows one after the other for $\tau_i^0 + g_i^{\max}$ seconds each. After the traffic situation has relaxed, i.e. as soon as all queues can be cleared again within the desired service interval T , the control is handed over to the optimization strategy. This uses the available free intersection capacity T^{res} according to equation (26) for flexible switching sequences or green time extensions, i.e. for more frequent set-ups or idling periods, as long as it helps to save waiting times. Such a scenario is illustrated in figure 8: at an initially over-saturated intersection, the stabilization strategy manages to reduce the queue lengths, before it hands over to the optimization strategy, which lets the queue lengths exponentially converge to the optimum cycle associated with minimum waiting times.

5. Simulation of the self-organized traffic light control

We have simulated the above control strategy (17) with the macroscopic network flow model sketched in section 2, using our short-term flow anticipation algorithm (see section 3). For comparison, we have also implemented our new control strategy in the commercial simulation tool VISSIM [105], which uses a microscopic car-following model and represents the dynamics of vehicles in more detail. This has resulted in qualitatively the same and quantitatively very similar results, so that we do not show these duplicating results here.

For simplicity, our computer simulations assume that all traffic flows at the intersections are incompatible, i.e. only one traffic flow can be served at a time. In the following, we will report the corresponding simulation results and analyze the performance of our control strategy.

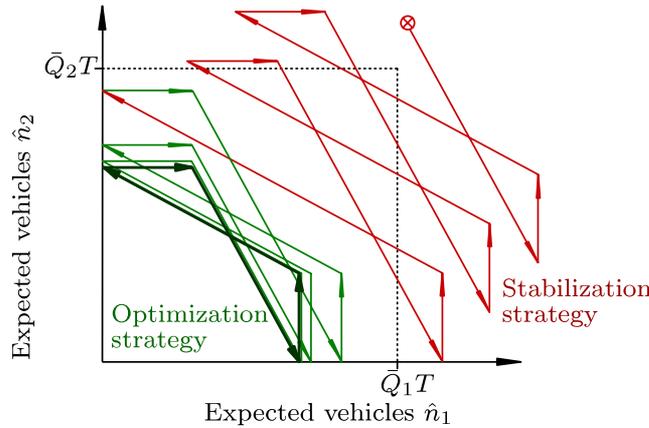


Figure 8. Mutual, time-dependent interdependence of the expected number of vehicles (\hat{n}_1, \hat{n}_2) at an intersection with two identical traffic flows. The initial state (crossed circle) corresponds to an over-saturated traffic condition. To clear the queues as fast as possible, the stabilization strategy (red) minimizes switching losses by serving each traffic flow exactly once within the desired service interval T . As soon as the trajectories are below the critical threshold $n_i^{\text{crit}} = \bar{Q}_i T$ defined by equation (20) in the limit $T^{\text{max}} \rightarrow \infty$, the optimization strategy is being activated (green lines). The optimization strategy uses the available free intersection capacity to converge towards the fastest possible switching sequence, which is the optimum traffic light cycle in terms of travel time minimization.

5.1. Operation modes at an isolated intersection

As a first test scenario, we study an isolated intersection with four traffic flows as depicted in figure 9. We are interested in the average total queue length $\bar{n} = \langle \sum_i n_i \rangle$ in the steady state, i.e. over one simulation hour. Whereas the inflow on the side streets was set to a constant volume of $Q_B = Q_D = 180$ vehicles per hour, the inflow $Q_A = Q_C$ on the two-lane main streets was varied. With a saturation flow rate of 1800 vehicles per hour and lane, we had $Q_A^{\text{max}} = Q_C^{\text{max}} = 3600$ vehicles per hour and $Q_B^{\text{max}} = Q_D^{\text{max}} = 1800$ vehicles per hour. Furthermore, the set-up times to switch between traffic flows were $\tau_i^0 = 5$ s. With the control parameters $T = 120$ s and $T^{\text{max}} = 180$ s for the desired and the maximum service intervals, respectively, the sufficient stability condition equation (24) was satisfied, if the utilization

$$u = \sum_i Q_i / Q_i^{\text{max}} \quad (27)$$

was less than 0.83. This means that our traffic light control was stable as long as the average inflow on the main streets $Q_A = Q_C$ was less than 1140 vehicles per hour.

For different levels of saturation, our self-organized traffic light control exhibits several distinct operation regimes:

5.1.1. Serving single vehicles at low utilizations. In the low-utilization regime, traffic demand is considerably below capacity. A minimization of the *average* waiting times

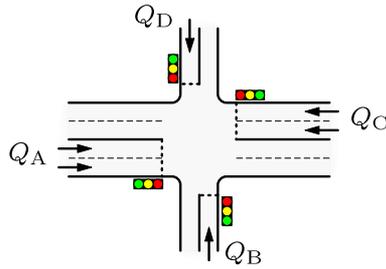


Figure 9. Isolated intersection with four competing traffic flows.

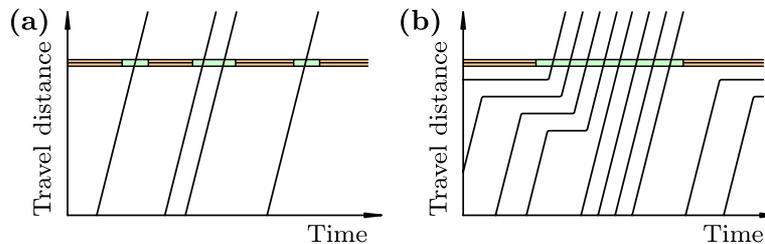


Figure 10. Illustration of vehicle trajectories for different operation modes: (a) in the low-utilization regime, the vehicles are served by a green light just upon their arrival at the stop-line (horizontal bar). Thereby, a stopping of the vehicles can be avoided. (b) At higher utilizations, the formation of vehicle platoons is unavoidable. However, serving vehicle platoons rather than maintaining the first-come-first-served principle allows one to minimize the average waiting time, as switching losses are reduced.

is achieved by serving the vehicles just upon their arrival, i.e. according to a first-come-first-served principle. This operation mode, which also minimizes *individual* travel times, is illustrated in figure 10(a).

5.1.2. Service of platoons at moderate utilizations. As the traffic demand increases, several vehicles may arrive at the intersection at about the same time, i.e. they may mutually obstruct each other. Some vehicles will have to wait, which implies the formation of platoons. However, given a certain utilization level, serving platoons becomes more efficient than applying the first-come-first-served principle (see figure 10(b)): the reduction of switching losses by serving platoons rather than single vehicles does not only reach a higher intersection capacity, but also a minimization of the average travel times.

5.1.3. Suppression of minor flows at medium utilizations. In figure 11, for utilizations u between about 0.3 and 0.5 one can see that the (multi-lane) main streets are served more frequently than the (one-lane) side streets. That means the interruption of the main flows by minor flows is suppressed, which is again in favor of minimizing the average waiting times.

5.1.4. Flow stabilization at high utilizations. For even higher utilizations u , our self-organized traffic light control does not exclusively follow the travel time optimization

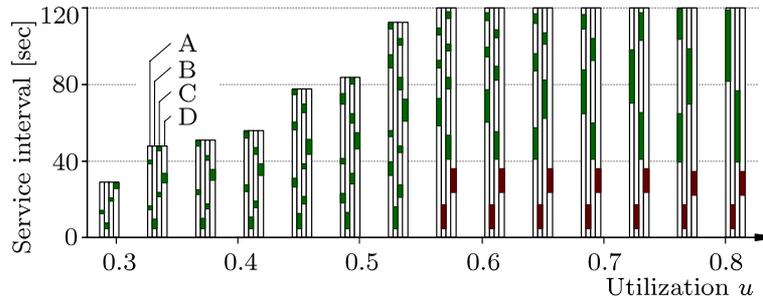


Figure 11. Flexible switching sequences for different utilization levels u over a complete service interval of traffic flow B. The results were obtained by computer simulation of the traffic flows at a single intersection after a transient time period. For details see section 5.2.

strategy any longer: the side streets would be served too rarely or too briefly. This becomes clear in figure 11 for utilizations u above 0.55. An efficient usage of the intersection capacity is now reached by serving the side streets as soon as their vehicle queues have reached a critical size. Thereby, the stabilization mechanism (see section 4.4) ensures that the safety-critical service interval of $T = 120$ s is never exceeded. Interestingly, there have emerged switching sequences of higher periods, that is, it may require several service intervals T before a switching sequence repeats. Nevertheless, all remaining capacity is still used to serve the main streets in the most flexible way, i.e. by serving them as often as possible.

5.2. Performance at an isolated intersection

Let us now compare our self-organizing control strategy with a simple fixed-time cycle-based strategy, the cycle time of which was set to a constant value of 120 s. While the switching order was set to A–B–C–D, the green times were adapted according to the formula

$$g_i^0 = \frac{u_i}{\sum_{i'} u_{i'}} \left(T - \sum_{i'} \tau_{i'}^0 \right). \quad (28)$$

That is, the green time g_i^0 of each flow i was specified proportionally to the corresponding partial utilization $u_i = Q_i/Q_i^{\max}$.

5.2.1. Constant inflows. Figure 12 shows the average total queue lengths \bar{n} for the case of regular inflows, i.e. for identical time gaps between the arriving vehicles. Interestingly, the optimization strategy performs better than the cycle-based approach as long as the traffic demand is low. But it fails at high utilizations $u > 0.6$, which is due to the strong prioritization of the main streets, where a higher throughput can be reached over a short optimization horizon. In the course of time, the side streets are, therefore, served too seldom or too briefly.

The stabilization strategy of section 4.4, in contrast, is stable at all utilization levels u , but it is associated with longer queues and, thus, longer average waiting times. However, the combined strategy of section 4.5 starts serving the side streets already *before* their

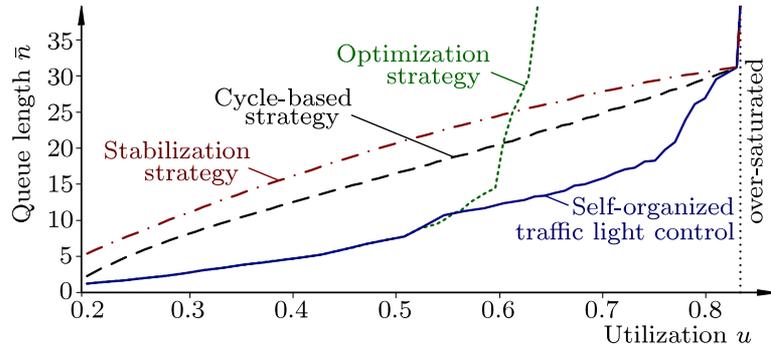


Figure 12. Average total queue length $\bar{n} = \sum_i \bar{n}_i$ at an intersection as depicted in figure 9. The optimization strategy becomes unstable already at medium utilization levels $u > 0.6$ and the stabilization strategy performs always worse than the cycle-based control. By suitably combining both inferior strategies, however, our self-organized traffic light control performs significantly better at *all* utilization levels. This also means that the traffic flow network will enter the over-saturated flow regime later, if at all. Therefore, traffic breakdowns during rush hours can be avoided or at least delayed, and the recovery from congestion will proceed faster.

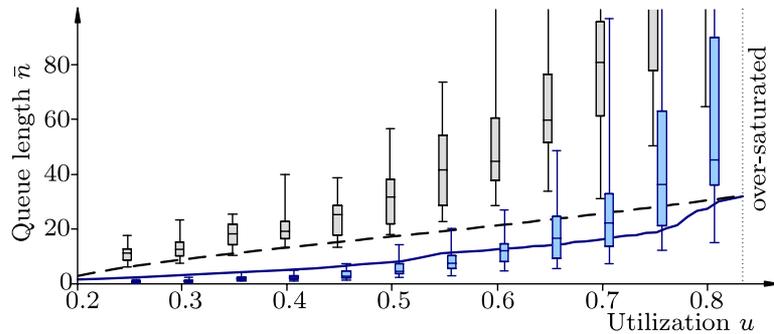


Figure 13. Box-whisker plot for the queue lengths at an isolated intersection (see figure 9) now with stochastic inflows. For the cycle-based control, the queue lengths are significantly higher compared to the case with regular inflow (see figure 12). In contrast, our self-organized control strategy manages to adjust to the stochastic variations in a flexible way, which leads to a reduction in both the mean value and the variance of the queue lengths.

queues grow too long. For this reason, the corresponding self-organized control strategy reaches a significant reduction of queue lengths and waiting times at *all* utilization levels.

5.2.2. Variable inflows. In the following simulation we assume that the vehicles arrive in platoons, where both the size of the platoons as well as the time gap between them are Poisson-distributed. Figure 13 shows the box-whisker plot (0–25–50–75–100 percentiles) of the stationary queue length distribution over 25 independent simulation runs each.

Because the cycle-based control strategy cannot respond to irregular inflow patterns, the green times are sometimes too short and sometimes too long, resulting in greater

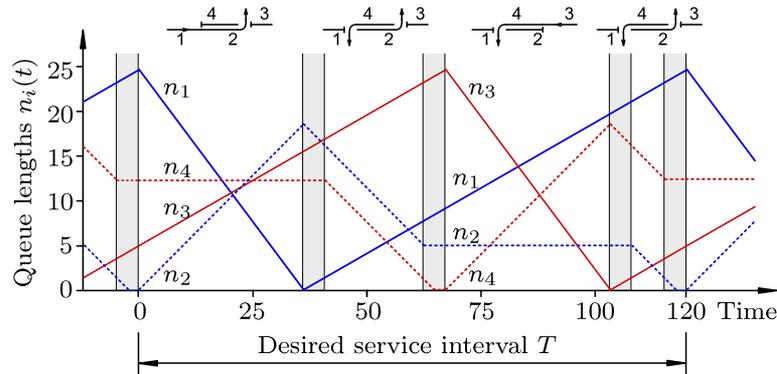


Figure 14. Time-dependent queue lengths for the non-acyclic road network of figure 1(b) analyzed in figure 2, but now assuming our self-organized traffic light control. We find stable queue lengths $n_i(t)$ for the same parameters and boundary conditions that caused a dynamic instability when a clearing policy was applied (see figure 2). The inflows were $Q_A = Q_B = 1100$ vehicles per hour, the saturation flow rates were 1800 vehicles per hour and lane, and the set-up times (vertical bars) were $\tau_i^0 = 5$ s. Our self-organized traffic light control stabilizes the network and serves each traffic flow once within the desired service interval $T = 120$.

delays at all utilization levels. In contrast, the self-organizing traffic light control has a large degree of flexibility to adjust to randomly arriving platoons. At low utilizations $u < 0.5$, where it is possible to serve the platoons just as they arrive, there are almost no delays. But even at higher utilizations $u \leq 0.7$, the queue lengths are significantly smaller compared to the case with regular inflows (see figure 13). Hence, our self-optimizing traffic lights could adjust well to the fluctuations in the inflow: the irregularly arriving platoons were served by irregular switching sequences. Altogether, this resulted in a reduction of the variability of the queue lengths and the related waiting times.

5.3. Coordination in networks

5.3.1. Solving the Kumar–Seidman problem. In appendix A.1, we demonstrate how a clearing policy, e.g. the Clear-Largest-Buffer-Strategy, can behave unstably in non-acyclic networks. The same network, illustrated in figure 1(b), shall now be operated with our self-organized traffic light control. Figure 14 shows the periodic behavior of the queue lengths in the steady state. Our self-control succeeds to stabilize the network, in particular because the stabilization strategy terminates serving streets 2 and 4 as soon as the anticipated number of vehicles on streets 3 or 1, respectively, has exceeded the critical threshold $n_i^{\text{crit}}(z_i)$ given by equation (20). Inefficiencies due to overly long green time extensions, which were responsible for the instability when using clearing policies, are thereby avoided.

5.3.2. Irregular networks. Let us now consider a 9×9 lattice road network, where both the length and the number of lanes of the road sections are irregular. The network layout is depicted in figure 15(a). The saturation flow is 1800 vehicles per hour and lane, and the speed limit is 50 km h^{-1} on all streets. Traffic enters and leaves the network at its boundary links and distributes according to a constant turning fraction of $\alpha_{ij} = 10\%$ turning left and right, while $\alpha_{ij} = 80\%$ go straight ahead at each intersection. The arrival

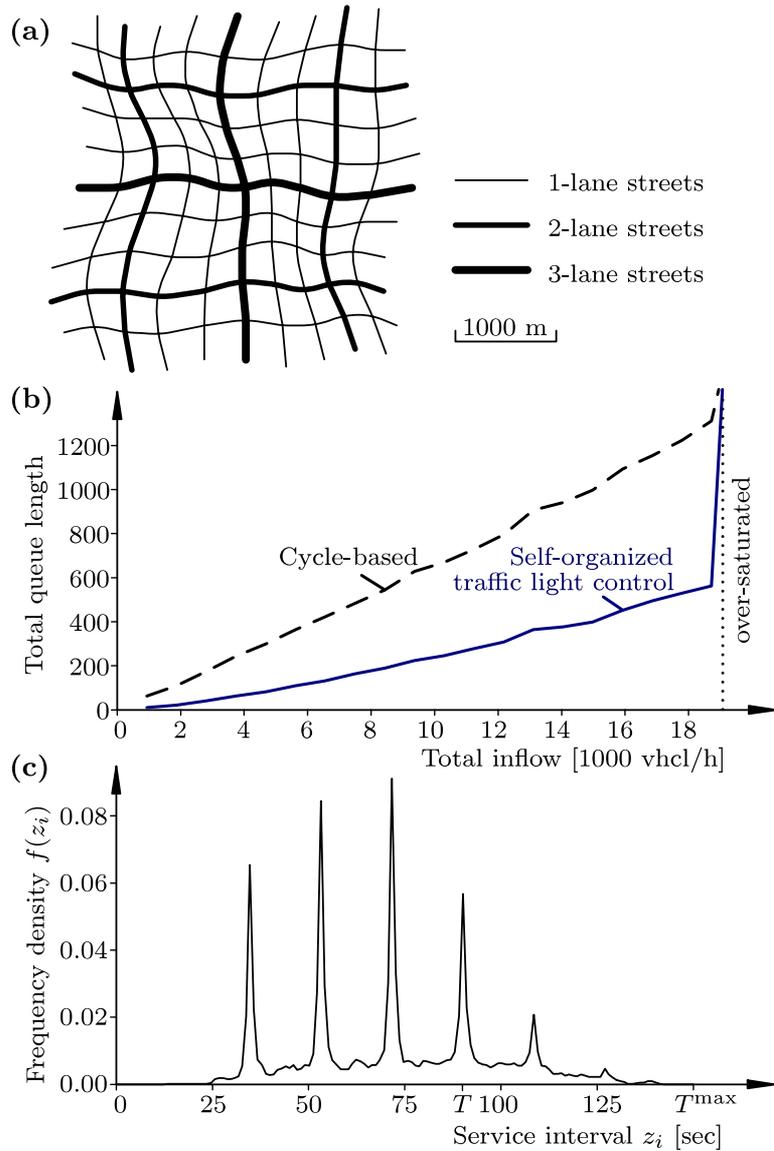


Figure 15. (a) Road network with irregular road lengths and different numbers of lanes. (b) Average queue lengths for different inflow rates. In contrast to the cycle-based strategy, our approach reaches a substantial reduction in the average queue length and related travel times. (c) The frequency density distribution $f(z_i)$ exhibits prominent peaks at different service intervals z_i . This indicates a self-organized coordination with a tendency of cycle times that are multiples of ≈ 18 s. Interestingly enough, a cycle of 18 s duration does not occur itself.

rate at each entry point is proportional to the corresponding number of lanes. For the operation of the traffic lights we assume a set-up time of $\tau_i^0 = 5$ s, a desired service interval of $T = 90$ s and a maximum service interval of $T^{\max} = 150$ s. For the fixed-time control strategy, with which we compare our results, we chose a cycle time of 90 s, demand-adaptive green times as specified by equation (28) and random offsets between the intersections.

Figure 15(b) plots the total average queue length $\sum_i \bar{n}_i$ in the stationary state (over one simulation hour) against the total traffic volume entering the network. Above a maximum inflow of 18 700 vehicles per hour, where the first intersections are over-saturated, neither strategy can prevent the queues from growing. Up to this value, however, our self-organized control strategy exhibits significantly smaller queue lengths in contrast to the cycle-based strategy. This is particularly due to the fact that our strategy has the flexibility to switch more often at less saturated intersections. However, the traffic lights are still coordinated and the gain in performance is significant. It is even higher than in the case of a single intersection.

The implicit coordination of the traffic lights becomes clear in figure 15(c), which has been determined for a total inflow of 10 000 vehicles per hour. It shows the frequency density distribution $f(z_i)$ of the service intervals z_i over all traffic lights in the network exhibits prominent peaks at fractions of a basic frequency. This regularity indicates that a distinct periodicity in the switching sequence has emerged. Even though many traffic flows are served exactly once within T , the period of the actual switching sequences is much smaller. This is because some traffic flows are served several times within the time period T . Therefore, it may take several intervals T before a switching sequence repeats (see also figure 11). Nevertheless, the service interval does not exceed the maximum service interval T^{\max} .

6. Conclusions and outlook

In this paper, we have proposed a self-organized traffic light control based on decentralized, local interactions. A visualization of its functional principles and properties is provided on the webpage <http://traffic.stefanlaemmer.de>, which includes many video animations of traffic simulations. The corresponding self-control concept is based on equation (17), together with the specifications in equations (16), (20) and (25). It differs from previous signal control approaches in the following points:

- (i) It reaches a superior performance by a non-periodic service, which is more flexible. A periodic traffic light control may, nevertheless, emerge, if the street network is grid-like and the incoming flows and turning fractions (or the boundary conditions) are periodic.
- (ii) The variation of waiting times is surprisingly small, i.e. the average waiting times are well predictable, even though the sequence and duration of green times are basically unpredictable.
- (iii) Our simulation results suggest that a substantial reduction of the average travel times, and therefore also of the fuel consumption and CO₂ emission, could be reached [35].
- (iv) The greatest gain in performance compared to previous traffic control approaches is expected (a) for strongly varying inflows, (b) irregular road networks, (c) large variations of the flows in different directions and among neighboring traffic lights, or (d) at night, where single vehicles should be served upon their arrival at the traffic light.

The success principle behind the superior performance of our decentralized self-control concept is the *combination* of two inferior strategies, a stabilization and an optimizing rule,

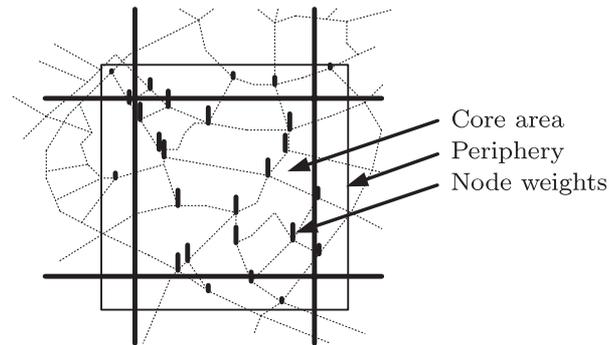


Figure 16. Illustration of the subdivision of a road network into non-overlapping subnetworks ('core areas') and definition of peripheral boundary areas as proposed in [106].

which allows for a varying sequence of traffic phases and a spatially coordinated, non-cyclical operation. The new approach can be easily integrated into a given traffic control environment (i.e. it is compatible with pre-specified controls at certain intersections). The decentralized sensor, communication and control concept is potentially less costly than centralized control concepts, and it can be set up in a way ensuring that traffic lights are still operational when measurement sensors or communication fail. Extensions to multi-phase operation and prioritization of public traffic will be presented in a forthcoming paper. Similar decentralized self-control strategies could be applied to the coordination of logistic or production processes and even to the coordination of work-flows in companies and administrations. (In fact, it might be easier to implement them in one of *these* systems, as traffic legislation and/or operation would first have to be adjusted to allow for the more efficient, non-cyclical traffic light operation.)

The proposed self-organized traffic light control is the first concrete realization of the approach suggested in a previous patent [106]. There, the road network has been completely subdivided into non-overlapping subnetworks ('core areas'). Moreover, each of the subnetworks was extended by additional neighboring nodes that define a boundary area ('periphery'). The boundary areas overlap with parts of the neighboring core areas and serve the coordination between the traffic light controls of the subnetworks (see figure 16). In each core area, one first determines highly performing solutions, assuming given traffic flows in the boundary areas. A traffic light control for the full network is then defined by a combination of highly performing traffic light controls for the subnetworks. The combination which performs best in the full network is finally applied (where the best solution for the full network is not necessarily the combination of the best solutions for the subnetworks).

The realization proposed in this paper assumes the smallest possible specification of the subnetworks, namely the single nodes and the corresponding set of ingoing links. The neighboring nodes constitute the respective boundary area of a node. The boundary areas are involved in the short-term anticipation of traffic flows in the associated subnetworks. To determine highly performing traffic light controls in the subnetworks, we apply optimization and stabilization strategies (see sections 4.3 and 4.4). The traffic light control in the full network is then implemented as follows: the stabilization strategy is applied at

nodes i where the set Ω_i is non-empty, while the optimization strategy is applied at the other nodes. This traffic light control performs better in the full network than applying the optimization strategy at all nodes, although the latter would minimize the travel times locally. The higher performance results from avoiding spill-back effects (see section 4.5), which would eventually block other intersections.

Acknowledgments

The authors are grateful to the German Research Foundation (DFG research projects He 2789/5-1, 8-1) and the Volkswagen Foundation (project I/82 697) for partial financial support of this research.

Appendix. General problems of network flow coordination

A.1. Dynamic instabilities

In game theory, it is known that the interaction of selfish agents can lead to inefficiencies, such as social dilemmas [107]. Therefore, a decentralized flow optimization by each single intersection is not necessarily optimal for the network. In fact, it is not a successful strategy to design a control algorithm for a single intersection and to operate the network based on such local intersection controls. Even if the control at each intersection minimizes the local increase of travel time, the dynamic coupling of neighboring intersections in the network can lead to inefficiencies due to correlations in the flow dynamics (see figure 2). The problem is either the loss of service times by frequent switches of traffic lights or the lack of coordination between them, which may increase the average waiting times. In other words, the intersections may not be able to handle the same amount of traffic as they could in isolation, assuming that the arrival of vehicles is continuous. Inefficiencies reduce the intersection capacities and cause the queues to grow longer and longer. The related spill-back effect sooner or later blocks the flow at upstream intersections. This phenomenon, which is referred to as dynamic instability [59], was, for example, demonstrated to occur under the following two conditions:

- (i) if the road network is non-acyclic and
- (ii) if the traffic control pursues a clearing policy.

In the following, both mechanisms shall be explained in more detail.

Regarding (i). A flow network is acyclic, if one could rank the nodes in such a way that all flows pass the nodes from lower to higher rank. Road networks, however, can never be acyclic. This is simply due to the fact that there always exist paths leading from one intersection to any other and back (not necessarily along the same route). This makes such a ranking impossible. The critical aspect of non-acyclic networks is that information propagates in so-called feedback loops. It means that, if one intersection sends a platoon of vehicles to one of its neighbors, it influences the time point at which another platoon is sent back. Thus, the arrivals at an intersection are not independent of its past switching sequence. Because these couplings do not only exist between neighboring intersections, but between all intersections in the network, and because these couplings have dynamically varying travel-time-related time delays, these feedback are far too complex to be anticipated locally.

Regarding (ii). Clearing policies continue serving a street until its queue has been fully cleared [57]. They only differ in the rules selecting what street to serve next. Such policies were shown in many experiments to be optimal at isolated intersections [81, 56, 108], but also to cause dynamic instabilities in non-acyclic networks [59, 58]. This is, for example, the case in the road network depicted in figure 1(b). Even though each intersection alone would have been controlled optimally, as soon as they are placed next to each other in a network, they turn out to behave unstably (see figure 2). These facts indicate the importance to test decentralized traffic light controls in non-acyclic networks. Unfortunately, many recently proposed approaches have been tested either at isolated intersections or in networks with unidirectional streets only. This may explain why most decentralized control concepts have not been practically implemented.

Our arguments above, however, do not imply that a road network can *never* be successfully controlled in a decentralized way, i.e. with independent control algorithms at each intersection. As shown in this paper, such a strategy is in fact possible, but it requires using a novel control mechanism. In section 4, we have proposed such a mechanism, designed a self-organizing traffic light control and extended it to fulfill critical safety requirements, i.e. to comply with maximum red times.

A.2. Chaotic dynamics

Switched flow networks are known to exhibit chaotic behavior under certain conditions [109, 65]. In principle, this is also expected to apply to traffic light controlled road networks. As a generic feature, chaotic behavior of a dynamical system is characterized by an exponential divergence of initially close trajectories. This sensitivity against small perturbations, which is a result of the intrinsic nonlinearity of the system, is often named the ‘butterfly effect’ [110]. It may occur even without any stochasticity in the system behavior.

Chase *et al* [47, 111] illustrated that chaotic behavior emerges even in very simple switched flow systems. For example, in the case of a single server responsible for serving three or more different flow directions, the resulting dynamics may be chaotic if the server is filling one buffer up to a certain level and then switches to another buffer (switched arrival system) [112]. But also the opposite case, where the server starts clearing a buffer as soon as its fill level exceeds a critical threshold, exhibits chaotic behavior (switched server system with limited buffers) [113, 114]. The latter case directly corresponds to a traffic light controlled intersection with restricted queue lengths at the incoming links. The generic mechanism leading to this behavior can be understood by studying the manifold (hyperplane), in which the trajectories of the related queue lengths (reflecting buffer fill levels) evolve. Because the underlying switching rules impose certain boundaries on this hyperplane, the trajectories experience a so-called ‘strange reflection’ if they hit one of these boundaries. This observation allows us to describe such systems in terms of ‘pseudo billiard dynamics’ [115, 116].

Studying the temporal evolution of vehicle positions, chaos can be observed even if the switching sequence of the traffic lights is given. This was shown by Toledo *et al* [117] and Nagatani [118] for a single vehicle moving through a sequence of fixed-time controlled traffic lights. This observation is independent of whether the distances between the traffic lights are regular or not [119]. In order to observe chaos, moreover, one does not even

require traffic lights at all. Wastavino *et al* [24] illustrated this for the case where vehicles are obstructed by yield signs.

The above examples suggest that chaotic behavior is intrinsic to vehicular flow in traffic networks. Whereas traffic flow is statistically well predictable at an aggregate level, it becomes highly unpredictable as soon as we want to describe its dynamics. Predictability, however, is of great importance for the design of a traffic light control that should be able to coordinate traffic flows, in particular to respond to large platoons as well as single vehicles. The purpose of our proposed anticipation strategy (see section 3), therefore, is not to statistically average over the complex nonlinear dynamics, but to cope with it on a short timescale in the most flexible way.

A.3. Limited prognosis time horizon

The unpredictable nature of traffic flow makes it particularly difficult to anticipate traffic conditions over long time horizons. Even if we assume the streets to be equipped with detectors and the intersections to communicate with each other, the prognosis time horizon can hardly be larger than twice the travel time L_i/V_i along the connecting links, e.g. 30–40 s for a typical road section with $L_i \approx 300$ m and $V_i \approx 50$ km h⁻¹. Whereas the model presented in section 2 can predict well over time horizons of less than the travel time L_i/V_i , larger horizons obviously need to take the switching sequence of neighboring intersections into account. If the control decision of an intersection depends on a time horizon of more than twice the travel time, this implies that the outcome of the control decision must be already known to its neighbor. Such kinds of information loops are yet another complication by the non-acyclic nature of road networks.

These considerations show that the problem of limited prognosis horizons is common to all flexible, vehicle-responsive traffic light controls, no matter whether they are implemented in a centralized or decentralized way. Thus, the fact that long-range interactions are highly complex and almost impossible to predict, holds for any control as long as it is flexibly responding to changing traffic conditions and not just imposing a pre-defined pattern on the traffic flows, such as conventional, cycle-based controls do. Another consequence of the limited prognosis time horizon is that any optimization is inevitably short-sighted and, therefore, must be regarded as a potential source of inefficiency and instability. This problem can be overcome, however, by introducing an appropriate stabilization strategy as presented in section 4.4.

References

- [1] Helbing D, 1997 *Verkehrsdynamik. Neue physikalische Modellierungskonzepte* (Berlin: Springer)
- [2] Chowdhury D, Santen L and Schadschneider A, *Statistical physics of vehicular traffic and some related systems*, 2000 *Phys. Rep.* **329** 199
- [3] Schadschneider A, *Traffic flow: a statistical physics point of view*, 2002 *Phys. Stat. Mech. Appl.* **313** 153
- [4] Helbing D, *Traffic and related self-driven many-particle systems*, 2001 *Rev. Mod. Phys.* **73** 1067
- [5] Helbing D and Molnár P, *Social force model for pedestrian dynamics*, 1995 *Phys. Rev. E* **51** 4282
- [6] Helbing D, Farkas I and Vicsek T, *Simulating dynamical features of escape panic*, 2000 *Nature* **407** 487
- [7] Helbing D, Johansson A, Mathiesen J, Jensen M H and Hansen A, *Analytical approach to continuous and intermittent bottleneck flows*, 2006 *Phys. Rev. Lett.* **97** 168001
- [8] Nagatani T, *The physics of traffic jams*, 2002 *Rep. Prog. Phys.* **65** 1331
- [9] Kerner B S, 2004 *The Physics of Traffic: Empirical Freeway Pattern Features, Engineering Applications, and Theory* (Berlin: Springer)
- [10] Nagatani T, *Chaotic jam and phase transition in traffic flow with passing*, 1999 *Phys. Rev. E* **60** 1535

- [11] Nishinari K, Treiber M and Helbing D, *Interpreting the wide scattering of synchronized traffic data by time gap statistics*, 2003 *Phys. Rev. E* **68** 067101
- [12] Kerner B S, *Empirical macroscopic features of spatial-temporal traffic patterns at highway bottlenecks*, 2002 *Phys. Rev. E* **65** 046138
- [13] Banks J H, *Investigation of some characteristics of congested flow*, 1999 *Transp. Res. Rec.* **1678** 128
- [14] Lighthill M and Whitham G, *On kinematic waves: ii. a theory of traffic flow on long crowded roads*, 1955 *Proc. R. Soc. A* **229** 317
- [15] Treiber M, Hennecke A and Helbing D, *Microscopic simulation of congested traffic*, 2000 *Traffic and Granular Flow '99: Social, Traffic, and Granular Dynamics* ed D Helbing, H J Herrmann, M Schreckenberg and D E Wolf (Berlin: Springer) pp 365–76
- [16] Nagel K and Schreckenberg M, *A cellular automaton model for freeway traffic*, 1992 *J. Physique* **2** 2221
- [17] Daganzo C F, *The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory*, 1994 *Transp. Res. B* **28** 269
- [18] Helbing D, *A section-based queueing-theoretical traffic model for congestion and travel time analysis in networks*, 2003 *J. Phys. A: Math. Gen.* **36** L593
- [19] Daganzo C F, *The cell transmission model. II: network traffic*, 1995 *Transp. Res. B* **29** 79
- [20] Esser J and Schreckenberg M, *Microscopic simulation of urban traffic based on cellular automata*, 1997 *Int. J. Mod. Phys. B* **8** 1025
- [21] Helbing D, Lämmer S and Lebacque J-P, *Self-organized control of irregular or perturbed network traffic*, 2005 *Optimal Control and Dynamic Games* ed C Deissenberg and R F Hartl (Dordrecht: Springer) pp 239–74
- [22] Helbing D, Jiang R and Treiber M, *Analytical investigation of oscillations in intersecting flows of pedestrian and vehicle traffic*, 2005 *Phys. Rev. E* **72** 046130
- [23] Helbing D, Siegmeier J and Lämmer S, *Self-organized network flows*, 2007 *Networks and Heterogeneous Media* **2** 193
- [24] Wastavino L A, Toledo B A, Rogan J, Zarama R, noza V M and Valdivia J A, *Modeling traffic on crossroads*, 2007 *Phys. Stat. Mech. Appl.* **381** 411
- [25] Gugat M, Herty M, Klar A and Leugering G, *Optimal control for traffic flow networks*, 2005 *J. Optim. Theor. Appl.* **126** 589
- [26] Daganzo C F, *Urban gridlock: macroscopic modeling and mitigation approaches*, 2007 *Transp. Res. B* **41** 49
- [27] Zheng J-F, Gao Z-Y and Zhao X-M, *Modeling cascading failures in congested traffic and transportation networks*, 2007 *Phys. Stat. Mech. Appl.* **385** 700
- [28] Simonsen I, Buzna L, Peters K, Bornholdt S and Helbing D, *Stationary network load models underestimate vulnerability to cascading failures*, 2007 *Preprint* 0704.1952
- [29] Brockfeld E, Barlovic R, Schadschneider A and Schreckenberg M, *Optimizing traffic lights in a cellular automaton model for city traffic*, 2001 *Phys. Rev. E* **64** 056132
- [30] Fouladvand M and Nematollahi M, *Optimization of green-times at an isolated urban crossroads*, 2001 *Eur. Phys. J. B* **22** 395
- [31] Chowdhury D and Schadschneider A, *Self-organization of traffic jams in cities: effects of stochastic dynamics and signal periods*, 1999 *Phys. Rev. E* **59** R1311
- [32] Huang D and Huang W, *Traffic signal synchronization*, 2003 *Phys. Rev. E* **67** 056124
- [33] Sasaki M and Nagatani T, *Transition and saturation of traffic flow controlled by traffic lights*, 2003 *Phys. Stat. Mech. Appl.* **325** 531
- [34] Sekiyama K, Nakanishi J, Takagawa I, Higashi T and Fukuda T, *Self-organizing control of urban traffic signal network*, 2001 *IEEE Int. Conf. Syst. Man. Cybern.* **4** 2481
- [35] Lämmer S, *Reglerentwurf zur dezentralen Online-Steuerung von Lichtsignalanlagen in Straßennetzwerken (Controller design for a decentralized control of traffic lights in urban road networks)*, 2007 PhD Thesis Dresden University of Technology
- [36] Fouladvand M E, Shaebani M R and Sadjadi Z, *simulation of intelligent controlling of traffic flow at a small city network*, 2004 *J. Phys. Soc. Japan* **73** 3209
- [37] Barlovic R, Huisinga T, Schadschneider A and Schreckenberg M, *Adaptive traffic light control in the ChSch model for city traffic*, 2004 *Traffic and Granular Flow'03* ed P H L Bovy, S P Hoogendoorn, M Schreckenberg and D E Wolf (Berlin: Springer)
- [38] Gershenson C, *Self-organizing traffic lights*, 2005 *Complex Syst.* **16** 29
- [39] Lämmer S, Kori H, Peters K and Helbing D, *Decentralised control of material or traffic flows in networks using phase-synchronisation*, 2006 *Phys. Stat. Mech. Appl.* **363** 39
- [40] Nakatsuji T, Seki S and Kaku T, *Development of a self-organizing traffic control system using neural network models*, 1995 *Transp. Res. Rec.* **1324** 137

- [41] Ledoux C, *An urban traffic flow model integrating neural networks*, 1997 *Transp. Res. C* **5** 287
- [42] Mikami S and Kakazu Y, *Genetic reinforcement learning for cooperative traffic signal control*, 1994 *IEEE Conf. Intell.* **1** 223
- [43] Chiu S and Chand S, *Self-organizing traffic control via fuzzy logic*, 1993 *Decis. Control* **2** 1897
- [44] Trabia M B, Kaseko M S and Ande M, *A two-stage fuzzy logic controller for traffic signals*, 1999 *Transp. Res. C* **7** 353
- [45] Hoar R, Penner J and Jacob C, *Evolutionary swarm traffic: if ant roads had traffic lights*, 2002 *Proc. Congr. Evol. Comput.* **2** 1910
- [46] Perkins J R, Humes C and Kumar P R, *Distributed scheduling of flexible manufacturing systems: stability and performance*, 1994 *IEEE Trans. Robot. Autom.* **10** 133
- [47] Chase C and Ramadge P J, *On real-time scheduling policies for flexible manufacturing systems*, 1992 *IEEE Trans. Autom. Control* **37** 491
- [48] Burgess K and Passino K M, *Stable scheduling policies for flexible manufacturing systems*, 1997 *IEEE Trans. Autom. Control* **42** 420
- [49] Righter R, *Scheduling in multiclass networks with deterministic service times*, 2002 *Queueing Syst.* **41** 305
- [50] Savkin A V and Evans R J, 2002 *Hybrid Dynamical Systems* (Boston, MA: Birkhäuser)
- [51] Lefeber E, *Nonlinear models for control of manufacturing systems*, 2004 *Nonlinear Dynamics of Production Systems* ed G Radons and R Neugebauer (New York: Wiley) pp 71–83
- [52] Lan W-M and Olsen T L, *Multiproduct systems with both setup times and costs: fluid bounds and schedules*, 2006 *Oper. Res.* **54** 505
- [53] Eekelen J A W M, *Modelling and control of discrete event manufacturing flow lines*, 2007 PhD Thesis Eindhoven University of Technology
- [54] de Schutter B and de Moor B, *The extended linear complementarity problem and the modeling and analysis of hybrid systems*, 1999 *Hybrid Systems V* (Springer Lecture Notes in Computer Science vol 1567) ed P Antsaklis, W Kohn, M Lemmon, A Nerode and S Sastry (Berlin: Springer) pp 70–85
- [55] de Schutter B, *Optimizing acyclic traffic signal switching sequences through an extended linear complementarity problem formulation*, 2002 *Eur. J. Oper. Res.* **139** 400
- [56] Lefeber E and Rooda J E, *Controller design for switched linear systems with setups*, 2006 *Phys. Stat. Mech. Appl.* **363** 48
- [57] Perkins J and Kumar P R, *Stable, distributed, real-time scheduling of flexible manufacturing/assembly/disassembly systems*, 1989 *IEEE Trans. Autom. Control* **34** 139
- [58] Kumar P and Meyn S P, *Stability of queueing networks and scheduling policies*, 1995 *IEEE Trans. Autom. Control.* **40** 251
- [59] Kumar P R and Seidman T I, *Dynamic instabilities and stabilization methods in distributed real-time scheduling of manufacturing systems*, 1990 *IEEE Trans. Autom. Control.* **35** 289
- [60] Humes C, *A regulator stabilization technique: Kumar–Seidman revisited*, 1994 *IEEE Trans. Autom. Control.* **39** 191
- [61] Reiman M I and Wein L M, *Dynamic scheduling of a two-class queue with setups*, 1998 *Oper. Res.* **35** 532
- [62] Dai J G and Jennings O B, *Stabilizing queueing networks with setups*, 2004 *Math. Oper. Res.* **29** 891
- [63] Savkin A V, *Controllability of complex switched server queueing networks modelled as hybrid dynamical systems*, 1998 *Decis. Control* **4** 4289
- [64] Lämmer S, Donner R and Helbing D, *Anticipative control of switched queueing systems*, 2007 *Eur. Phys. J. B* at press
- [65] Wiggins S, *Introduction to applied nonlinear dynamical systems and chaos*, 2003 *Texts in Applied Mathematics* (Berlin: Springer)
- [66] Janson B N, *Dynamic traffic assignment for urban road networks*, 1991 *Transp. Res. B* **25** 143
- [67] Bowman J L and Ben-Akiva M E, *Activity-based disaggregate travel demand model system with activity schedules*, 2001 *Transp. Res. Pol. Pract.* **35** 1
- [68] Lan C-J, *Adaptive turning flow estimation based on incomplete detector information for advanced traffic management*, 2001 *Intell. Transp. Syst.* (Piscataway, NJ: IEEE) pp 830–5
- [69] Daganzo C F, *Queue spillovers in transportation networks with a route choice*, 1998 *Transp. Sci.* **32** 3
- [70] Herty M and Klar A, *Simplified dynamics and optimization of large scale traffic networks*, 2004 *Math. Models Methods Appl. Sci.* **14** 579
- [71] Garavello M and Piccoli B, *Traffic Flow on a Road Network Using the AwRascle Model*, 2006 *Commun. Part. Differ. Equations* **31** 243
- [72] Troutbeck R J and Brilon W, *Unsignalized intersection theory*, 1997 *Traffic Flow Theory: A State-of-the-Art Report* ed N Gartner, H Mahmassani, C J Messer, H Lieu, R Cunard and A K Rathi (London: Transportation Research Board) pp 8.1–8.47

- [73] Webster F V, 1958 *Traffic Signal Settings Road Research Laboratory Technical Paper* vol 39, pp 1–44
- [74] Smith M, Clegg J and Yarrow R, *Modeling traffic signal control*, 2001 *Handbook of Transport Systems and Traffic Control* ed K J Button and D A Hensher (Oxford: Pergamon) chapter 34, pp 503–26
- [75] Arfken G B and Weber H-J, 1995 *Mathematical Methods for Physicists* 4 edn (London: Academic)
- [76] Papadimitriou C H and Tsitsiklis J N, *The complexity of optimal queuing network control*, 1999 *Math. Oper. Res.* **24** 293
- [77] Porche I, Sampath M, Sengupta R, Chen Y-L and Lafortune S, *A decentralized scheme for real-time optimization of traffic signals*, 1996 *Proc. IEEE Conf. Control Appl.* pp 582–9
- [78] Papageorgiou M, Diakaki C, Dinopoulou V and Wang A K Y, *Review of road traffic control strategies*, 2003 *Proc. IEEE* **91** 2043
- [79] McDonald M and Hounsell N B, *Road traffic control: transyt and scoot*, 1991 *Concise Encyclopedia of Traffic & Transportation Systems (Advances in Systems Control and Information Engineering)* ed M Papageorgiou (Oxford: Pergamon) p 400
- [80] Helbing D, Molnár P, Farkas I and Bolay K, *Self-organizing pedestrian movement*, 2001 *Environ. Plann. Plann. Des.* **28** 361
- [81] van Oyen M P, Pandelis D G and Teneketzis D, *Optimality of index policies for stochastic scheduling with switching penalties*, 1992 *J. Appl. Probab.* **29** 957
- [82] Rothkopf M H and Smith S A, *There are no undiscovered priority index sequencing rules for minimizing total delay costs*, 1984 *Oper. Res.* **32** 451
- [83] Serres Y D, *Simultaneous optimization of flow control and scheduling in a single server queue with two job classes*, 1991 *Oper. Res. Lett.* **10** 103
- [84] Panwalker S S and Iskander W, *A survey of scheduling rules*, 1977 *Oper. Res.* **25** 45
- [85] Kumar S and Kumar P R, *Performance bounds for queueing networks and scheduling policies*, 1994 *IEEE Trans. Autom. Control* **39** 1600
- [86] Harrison J M, *A priority queue with discounted linear costs*, 1975 *Oper. Res.* **23** 260
- [87] Balachandran K R, *Parametric priority rules: an approach to optimization in priority queues*, 1970 *Oper. Res.* **18** 526
- [88] Duenyas I and van Oyen M P, *Heuristic scheduling of parallel heterogeneous queues with set-ups*, 1996 *Manag. Sci.* **42** 814
- [89] Baras J S, Dorsey A J and Makowski A M, *Two competing queues with linear costs and geometric service requirements: the μ -rule is often optimal*, 1985 *Adv. Appl. Probab.* **17** 186
- [90] Gartner N H, Little J D C and Gabbay H, *Optimization of traffic signal settings by mixed-integer linear programming part I: the network coordination problem*, 1975 *Transp. Sci.* **9** 321
- [91] Gartner N H, Little J D C and Gabbay H, *Optimization of traffic signal settings by mixed-integer linear programming part II: the network synchronization problem*, 1975 *Transp. Sci.* **9** 344
- [92] Gershenson C, *Design and control of self-organizing systems*, 2007 PhD Thesis Center Leo Apostel for Interdisciplinary Studies, Vrije Universiteit Brussel
- [93] Chang K C and Sandhu D, *Mean waiting time approximations in cyclic-service systems with exhaustive limited service policy*, 1992 *Perform. Eval.* **15** 21
- [94] Chang K C and Sandhu D, *Delay analyses of token-passing protocols with limited token holding times*, 1994 *IEEE Trans. Commun.* **42** 2833
- [95] Ram R and Viswanadham N, *GSPN models for versatile multi-machine workcenters with finite buffers*, 1993 *IEEE Int. Conf. Syst. Man. Cybern.* **2** 186
- [96] Retting R A and Greene M A, *Influence of traffic signal timing on red-light running and potential vehicle conflicts at urban intersections*, 1997 *Transp. Res. Rec.* **1595** 1
- [97] Porter B E and England K J, *Predicting red-light running behavior: a traffic safety study in three urban settings*, 2000 *J. Saf. Res.* **31** 1
- [98] Datta T K, Feber D, Schattler K and Datta S, *Effective safety improvements through low-cost treatments*, 2000 *Transp. Res. Rec.* **1734** 1
- [99] Perry D, Stadje W and Zacks S, *Contributions to the theory of first-exit times of some compound processes in queueing theory*, 1999 *Queueing Syst.* **33** 369
- [100] Abdel-Hameed M, *Optimal control of a dam using $p_{\lambda, \tau}^n$ policies and penalty cost when the input process is a compound Poisson process with positive drift*, 2000 *J. Appl. Probab.* **37** 408
- [101] Savkin A V, *Regularizability of complex switched server queueing networks modelled as hybrid dynamical systems*, 1998 *Syst. Control Lett.* **35** 291
- [102] Savkin A V, *Optimal distributed real-time scheduling of flexible manufacturing networks modeled as hybrid dynamical systems*, 2003 *Proc. 42nd IEEE Conf. on Decision and Control* vol 5, p 5468
- [103] Gazis D C, *Optimum control of a system of oversaturated intersections*, 1964 *J. Oper. Res. Soc. Am.* **12** 815

- [104] Gazis D C, *Traffic theory*, 2002 *International Series in Operations Research & Management Science* (Berlin: Springer)
- [105] <http://www.ptv.de>
- [106] Helbing D and Lämmer S, *Verfahren zur Koordination konkurrierender Prozesse oder zur Steuerung des Transports von mobilen Einheiten innerhalb eines Netzwerkes (Method for coordination of concurrent processes for control of the transport of mobile units within a network)*, 2006 Patent WO/2006/122528
- [107] Huberman B A and Lukose R M, *Social dilemmas and internet congestion*, 1997 *Science* **277** 535
- [108] van Eekelen J A W M, Lefeber E and Rooda J E, *State feedback control of switching server flowline with setups*, 2007 *Proc. Am. Control Conf.* pp 3618–23
- [109] Safonov L A, Tomer E, Strygin V V, Ashkenazy Y and Havlin S, *Multifractal chaotic attractors in a system of delay-differential equations modeling road traffic*, 2002 *Chaos* **12** 1006
- [110] Schuster H G and Just W, 2005 *Deterministic Chaos* (New York: Wiley–VCH)
- [111] Chase C, Serrano J and Ramadge P J, *Periodicity and chaos from switched flow systems: contrasting examples of discretely controlled continuous systems*, 1993 *IEEE Trans. Autom. Control* **38** 70
- [112] Rem B and Armbruster D, *Control and synchronization in switched arrival systems*, 2003 *Chaos* **13** 128
- [113] Peters K, Worbs J, Parlitz U and Wiendahl H-P, *Manufacturing systems with restricted buffer sizes*, 2004 *Nonlinear Dynamics of Production Systems* ed G Radons and R Neugebauer (New York: Wiley) pp 39–54
- [114] Yu G-X and Vakili P, *Periodic and chaotic dynamics of a switched-server system undercorridor policies*, 1996 *IEEE Trans. Autom. Control* **41** 584
- [115] Blank M and Bunimovich L, *Switched flow systems: pseudo billiard dynamics*, 2004 *Dyn. Syst. Int. J.* **19** 359
- [116] Peters K and Parlitz U, *Hybrid systems forming strange billiards*, 2003 *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **19** 2575
- [117] Toledo B A, Munoz V, Rogan J, Tenreiro C and Valdivia J A, *Modeling traffic through a sequence of traffic lights*, 2004 *Phys. Rev. E* **70** 016107
- [118] Nagatani T, *Chaos and dynamical transition of a single vehicle induced by traffic light and speedup*, 2005 *Phys. Stat. Mech. Appl.* **348** 561
- [119] Nagatani T, *Control of vehicular traffic through a sequence of traffic lights positioned with disordered interval*, 2006 *Phys. Stat. Mech. Appl.* **368** 560