Roncoli, Claudio; Bekiaris-Liberis, Nikolaos; Papageorgiou, Markos

Lane-changing feedback control for efficient lane assignment at motorway bottlenecks

Published in:
Transportation Research Record

DOI:
10.3141/2625-03

Published: 01/01/2017

Document Version
Peer reviewed version

Please cite the original version:
LANE-CHANGING FEEDBACK CONTROL FOR EFFICIENT LANE ASSIGNMENT AT MOTORWAY BOTTLENECKS

Claudio Roncoli, Corresponding Author
Dynamic Systems and Simulation Laboratory
Technical University of Crete
Chania, 73100, Greece
Tel: +30 28210 37289; Fax: +30 28210 37584; Email: croncoli@dssl.tuc.gr

Nikolaos Bekiaris-Liberis
Dynamic Systems and Simulation Laboratory
Technical University of Crete
Chania, 73100, Greece
Email: nikos.bekiaris@gmail.com

Markos Papageorgiou
Dynamic Systems and Simulation Laboratory
Technical University of Crete
Chania, 73100, Greece
Email: markos@dssl.tuc.gr

Word count: 5260 words + 9 tables/figures x 250 words (each) = 7510 words

August 1, 2016
ABSTRACT

We propose a feedback control strategy for lane assignment at bottleneck locations, assuming that a percentage of vehicles, equipped with Vehicle Automation and Communication Systems (VACS), are capable of receiving and executing specific lane-changing orders or recommendations. Starting from a previously proposed optimal control strategy, based on a simplified multi-lane motorway traffic flow model and formulated as a linear quadratic regulator, we design a feedback control problem aiming at maximising the throughput at bottleneck locations while distributing, according to a given policy, the total density at the bottleneck area among the different lanes, via optimal lane assignment of vehicles upstream of the bottleneck. The feedback control decisions are based on real-time measurements of the traffic state and inflow. The proposed strategy is tested on a nonlinear first-order macroscopic multi-lane traffic flow model, which also accounts for the capacity drop phenomenon.

Keywords: Motorway traffic control, lane-changing control, connected/automated vehicles
In the near future, Vehicle Automation and Communication Systems (VACS) are expected to revolutionise the features and capabilities of individual vehicles. Among the wide range of potentially introduced VACS, some may be exploited to interfere with the driving behaviour via recommending, supporting, or even executing appropriately designed traffic control tasks, providing unprecedented opportunities to improve traffic control performance (1). On the other hand, the uncertainty regarding the future development of VACS calls for the design of control strategies that are robust with respect to the different types of these new systems, as well as to their penetration rate. A promising new feature that can be exploited for traffic management is lane-changing control.

The problem of modelling the distribution of vehicles among lanes, in case of ordinary traffic, has been addressed in a number of research works, including (2, 3, 4, 5, 6, 7, 8), which show that the lane distribution is affected, among others, by some characteristics of the network layout (e.g., the total number of lanes); however this choice is also behavioural, since every single driver may autonomously decide to stay in a slower lane accepting the lower speed, stay in the slower lane and overtake when necessary (for lower densities), or choosing to travel constantly in a faster lane (in higher densities). In addition, particularly at bottleneck locations (e.g., lane-drops, on-ramp merges), human drivers usually perform suboptimal lane-changes based on erroneous perceptions, which may trigger congestion, and, thus, deteriorate the overall travel time (9, 10). Last but not least, some of the mentioned empirical investigations indicate that, in conventional traffic, capacity flow is not reached simultaneously at all lanes, a feature that reduces the potentially achievable cross-lane capacity. We therefore envision that, in case a sufficient percentage of vehicles are equipped with VACS having vehicle-to-infrastructure (V2I) capabilities and appropriate lane-changing automatic controllers or advisory systems, the overall throughput at the bottleneck location may be improved by execution of specific lane-changing commands decided by a central decision maker.

The problem of assigning traffic flow among lanes for motorways under fully automated or semi-automated driving has been studied in numerous research works during the last decades. To tackle the high complexity of the problem, several assumptions are typically made, such as known and constant prevailing speeds along the motorway and absence of traffic congestion, thanks to the assumed (but not addressed) appropriate operation of other control actions (e.g., ramp metering) at the motorway entrances; also, structural assumptions are commonly considered in order to limit the (otherwise vast) space of potential path assignments. In his seminal work, Varaiya (11) proposed a hierarchical framework for a fully automated motorway, where the decisions on the lane-changing behaviour of vehicles are addressed within the link layer, which consists of a set of parallel decentralised link controllers, each of them addressing a corresponding motorway link (of about 2 km in length). Following this framework, several strategies have been proposed to solve the problem of lane assignment within the link layer, designing control methodologies suitable for real-time applications, including the definition of well-justified and structured heuristic rules (12); the implementation of lane routing algorithms (13); and the definition of control laws to stabilise traffic conditions (14). On the other hand, optimisation methods for path planning through lanes have been developed (15, 16, 17, 18), however the computation complexity of the proposed optimisation problems makes them hardly applicable in a real-time context. Lane-changing control has also been considered, together with variable speed limits and ramp metering, within integrated traffic management strategies (19, 20, 21).

Recently, a combined lane-changing and variable speed limits control strategy was devel-
oped by Zhang and Ioannou (10), with the purpose of avoiding lane-changes in the immediate proximity of a bottleneck, which, especially in the case of heavy vehicles, may lead to premature triggering of congestion. In particular, lane-changing commands delivered as recommendation to the drivers, are defined according to a set of case-specific rules. Furthermore, Guériau et al. (22) proposed a multi-agent decentralised framework with the aim of performing cooperative lane-changing tasks based on information exchange between vehicles and a road side unit located in the proximity of a bottleneck.

We recently proposed in (23) an optimal feedback control strategy, formulated as a linear quadratic regulator, where the solution is applied in the form of a linear state-feedback control law, which is highly efficient in real-time even for large-scale networks. The control strategy aims at regulating the lane assignment of vehicles upstream of a bottleneck location so as to maximise the bottleneck throughput, targeting critical densities at bottleneck locations as set-points. However, as a result, the traffic density distribution among different lanes may remain (roughly) constant under any demand scenario. Although this behaviour would not produce any negative impact on the traffic performance, it may be, in some circumstances, undesirable. As an example, one can imagine a two-lane motorway, where both lanes have the same characteristics (i.e., same critical densities): targeting critical densities as set-points would result in equal flows in both lanes for any traffic situation. This behaviour is not permitted, for example, in European motorways, where vehicles are obliged to travel in the rightmost (for right-hand traffic) available lane, while overtaking is only allowed on the left side. For North-American freeways this issue is less crucial since vehicle overtaking is allowed on any lane; however, also in this case, traffic authorities may, for various reasons, prefer different specific lane distributions. In order to incorporate this feature, we propose here a methodology that does not always aim at tracking the critical density but, through opportunely defined functions, it allows to distribute the total density at a bottleneck area, among the different lanes, according to a given policy.

In the remaining paper, we first present the control design framework for multi-lane motorways proposed in (23); we then reformulate the control problem and design a feedback control law in order to achieve different traffic density distribution for the various lanes at the bottleneck area. We then present simulation experiments, employing a first-order macroscopic traffic flow model featuring the capacity drop phenomenon, in order to evaluate the effectiveness of the developed methodology and to highlight the different traffic behaviour in terms of flow distribution; in a conclusive section, we highlight the main results of the paper and propose further research challenges.

LANE-CHANGING-BASED OPTIMAL CONTROL OF MULTI-LANE MOTORWAYS AT BOTTLENECKS

Bottlenecks in motorways

A motorway bottleneck is a location where the flow capacity upstream is higher than the flow capacity downstream of the bottleneck location. Bottleneck locations can be lane-drops, merge areas, zones with particular infrastructure layout (e.g., strong grade or curvature, tunnels) or with external capacity-reducing events (e.g. work-zones, incidents). The nominal bottleneck capacity is the maximum traffic flow that can be maintained at the bottleneck location if the traffic flow arriving from upstream is smaller than (or equal to) the bottleneck capacity. On the other hand, if the arriving flow is higher than the capacity, or the lane-changing behaviour leads to exceeding the capacity of at least one lane, the bottleneck is activated, generating a congestion starting at the bottleneck.
location and spilling-back for as long as the upstream arriving flow is sufficiently high. Empirical observations show that, whenever a bottleneck is activated, the maximum outflow that materialises (also called discharge flow) may be some 5 to 20 percent lower than the nominal bottleneck capacity, and the difference between these two values of flow is called capacity drop \((24, 25)\). To avoid or delay the activation of a bottleneck, and the related capacity drop phenomenon, various traffic control measures have been proposed and applied \((26)\). In this work, we assume that the proposed control strategy operates simultaneously with some other controller (e.g., ramp metering \((27)\) or mainstream traffic flow control \((28)\)) that guarantees that the flow approaching the bottleneck area does not exceed the overall capacity of the bottleneck and, therefore, assuming an appropriate operation of the proposed lane-changing controller, traffic congestion may be completely avoided.

### Linear multi-lane traffic flow model

We consider a multi-lane motorway that is subdivided into \(i = 0, \ldots, N\) segments of length \(L_i\). while each segment is composed of \(j = m_i, \ldots, M_i\) lanes, where \(m_i\) and \(M_i\) are the minimum and maximum indexes of lanes for segment \(i\). We denote each element of the resulting grid (see Figure 1) as a cell, which is indexed by \((i, j)\). The model is formulated in discrete time, considering the discrete time step \(T\), indexed by \(k = 0, 1, \ldots\), where the time is \(t = kT\). In order to account for any possible network topology, including lane-drops and lane-additions, both on the right and on the left sides of the motorway, we assume that \(j = 0\) corresponds to the segment(s) including the most right lane; consequently, \(m_i\) and \(M_i\) are defined as the minimum and maximum indexes \(j\), respectively, for which a lane exists within segment \(i\). For example, looking at the hypothetical motorway stretch depicted in Figure 1, \(m_0 = 0\) and \(M_0 = 4\), while \(m_3 = 1\) and \(M_3 = 3\). According to this definition, the total number of cells from the origin to segment \(i\) is \(H_i = \sum_{r=0}^{i} (M_r - m_r + 1)\), and the total number of cells for the whole stretch is \(\bar{H} = H_N\).

Each motorway cell \((i, j)\) is characterised by the traffic density \(\rho_{i,j}(k)\), defined as the number of vehicles present within the cell at time instant \(k\) divided by \(L_i\). Density dynamically evolves according to the following conservation law equation, see e.g. \((29)\),

\[
\rho_{i,j}(k+1) = \rho_{i,j}(k) + \frac{T}{L_i} [q_{i-1,j}(k) - q_{i,j}(k)] + \frac{T}{L_i} [f_{i,j-1}(k) - f_{i,j}(k)] + \frac{T}{L_i} d_{i,j}(k),
\]

where \(q_{i,j}(k)\) is the longitudinal flow leaving cell \((i,j)\) and entering cell \((i+1,j)\) during time interval \((k,k+1]\); \(f_{i,j}(k)\) is the net lateral flow moving from cell \((i,j)\) to cell \((i,j+1)\) during time interval \((k,k+1]\); and \(d_{i,j}(k)\) is the external flow entering the network in cell \((i,j)\), either from the mainstream or from an on-ramp, during time interval \((k,k+1]\). Depending on the network topology, some terms of Equation 1 may not be present. In particular, the inflow \(q_{i-1,j}(k)\) does not exist for the first segment of the network; the outflow \(q_{i,j}(k)\) does not exist for the last segment before a lane-drop; while lateral flow terms \(f_{i,j}(k)\) exist only for \(m_i \leq j < M_i\). Following previous considerations, the total number of lateral flow terms is \(\bar{F} = \bar{H} - N\).

In order to guarantee numerical stability (since the discrete-time system described by Equation 1 may come from a discretisation of a PDE \((30)\)), the time step \(T\) must respect the so-called CFL condition \((31)\):

\[
T \leq \min_{i,j} \frac{L_i}{v_{i,j}^{\text{max}}},
\]

where \(v_{i,j}^{\text{max}}\) is the maximum speed allowed in cell \((i,j)\).
Similar modelling approaches of multi-lane motorway traffic are considered also in (29, 32, 33). One aspect that is interesting to be pointed out is that the net lateral flow $f_{i,j}(k)$ is considered only in one direction, namely from right to left lanes; therefore, $f_{i,j}(k)$ is actually the difference between the flow leaving and entering lane $j$ at its left side. This simplification is useful for the subsequent control problem formulation, since lateral flows are treated as control inputs.

Let us consider the well-known relation

$$q_{i,j}(k) = \rho_{i,j}(k) v_{i,j}(k);$$

replacing Equation 3 into Equation 1 we obtain

$$\rho_{i,j}(k + 1) = \frac{T}{L_i} v_{i,j}(k) \rho_{i-1,j}(k) + \left[ 1 - \frac{T}{L_i} v_{i,j}(k) \right] \rho_{i,j}(k) + \frac{T}{L_i} \left[ f_{i-1,j}(k) - f_{i,j}(k) \right] + \frac{T}{L_i} d_{i,j}(k),$$

which, treating speeds $v_{i,j}(k)$ as known parameters, can be seen as a Linear Parameter Varying (LPV) system in the form

$$x(k + 1) = A(k)x(k) + Bu(k) + d(k)$$

where (time index $k$ is omitted to simplify notation)

$$x = [\rho_{0,m_0} \cdots \rho_{0,M_0} \rho_{1,m_1} \cdots \rho_{N,M_N}]^T \in \mathbb{R}^\bar{H},$$

$$u = [f_{0,m_0} \cdots f_{0,M_0} f_{1,m_0}(k) \cdots f_{N,M_N-1}]^T \in \mathbb{R}^\bar{F},$$

$$d = \left[ \frac{T}{L_0} d_{0,m_0} \cdots \frac{T}{L_0} d_{0,M_0} \frac{T}{L_1} d_{1,m_1} \cdots \frac{T}{L_N} d_{N,M_N} \right]^T \in \mathbb{R}^\bar{H}. $$

$A \in \mathbb{R}^{\bar{H} \times \bar{H}}$, composed of elements $a_{r,s}$, which represents the connection between pairs of subsequent cells connected by a longitudinal flow, and $B \in \mathbb{R}^{\bar{H} \times \bar{F}}$, composed of elements $b_{r,s}$, which
reflected the connection of adjacent cells connected by lateral flows, are defined as

\[ a_{r,s} = \begin{cases} 
1, & \text{if } r = s \text{ and } (j < m_{i+1} \text{ or } j > M_{i+1}) \\
1 - \frac{T}{L}, & \text{if } r = s \text{ and } (i = N \text{ or } m_{i+1} \leq j \leq M_{i+1}) \\
\frac{T}{L} v_{i-1,j}, & \text{if } r > H_0 \text{ and } s = r - M_{i-1} + m_i - 1 \\
0, & \text{otherwise}
\end{cases} \]  

\[ b_{r,s} = \begin{cases} 
\frac{T}{L}, & \text{if } j > m_i \text{ and } s = r - i \\
-\frac{T}{L}, & \text{if } j < M_i \text{ and } s = r - i + 1 \\
0, & \text{otherwise},
\end{cases} \]  

where \( r = \sum_{r=0}^{i-1} H_r + j - m_i \).

3 Optimal control problem formulation with constant set-points

The linear system described in the previous section is used for formulating an optimal control problem with the purpose of manipulating the lateral flows in order to avoid the creation of congestion due to the activation of a bottleneck. Under the assumption that the overall traffic flow entering the controlled area does not exceed significantly the bottleneck capacity and that the controller succeeds to avoid the creation of congestion, we can assume that the speeds in all cells remain at a constant value (e.g., the free flow speed) \( v_i,j(k) \equiv \bar{v}, \forall i, j, k \). In addition, we assume that the measurable inflows \( d \) are constant; note that actual slow time-variation of \( d \) will not affect the control performance significantly. With these assumptions, the system in Equation 5 can be viewed as a Linear Time Invariant (LTI) system

\[ x(k+1) = Ax(k) + Bu(k) + d. \]  

Identifying the nominal capacity of a bottleneck is a non trivial task; in fact, Elefteriadou et al. (34) and Lorenz and Elefteriadou (35) have demonstrated that the real flow capacity in a merge area may vary quite substantially from day to day even under similar environmental conditions; therefore, any control strategy attempting to achieve a pre-specified capacity flow value may either lead to overload and congestion (on days where the real capacity happens to be lower than its pre-specified target value) or to underutilisation of the infrastructure (on days where the real capacity happens to be higher than its pre-specified target value). On the other hand, the critical density, at which capacity flow occurs, exhibits smaller variations (36), and it is therefore preferable targeting a density set-point (i.e., the critical density) at the bottleneck location. In (23) we propose a control strategy that is always targeting the critical densities for each lane; and, for the case they are unknown, an extremum seeking algorithm (37) was proposed to estimate them.

We define the following quadratic cost function (over an infinite time horizon) that accounts for the penalisation of the difference between some (targeted) densities and the corresponding pre-specified (assumed constant) set-point values; as well as a penalty term aiming at maintaining small control inputs, i.e., small lateral flows (weighted by \( \phi \)):

\[ J = \sum_{k=0}^{\infty} \left\{ \sum_i \sum_j \alpha_{i,j} \left[ \rho_{i,j}(k) - \bar{\rho}_{i,j} \right]^2 + \phi \sum_i \sum_{j=m_i}^{M_i-1} [f_{i,j}(k)]^2 \right\}. \]  

TRB 2017 Annual Meeting Paper revised from original submittal.
where \((\hat{i}, \hat{j})\) denote the targeted cells, \(\hat{\rho}_{\hat{i}, \hat{j}}\) is the desired set-point, and \(\alpha_{\hat{i}, \hat{j}}\) is the corresponding weighting parameter. We rewrite Equation 12 in matrix form as

\[
J = \sum_{k=0}^{\infty} \left[ (C\hat{x}(k) - \hat{y})^T Q (C\hat{x}(k) - \hat{y}) + u^T(k) R u(k) \right],
\]

where \(Q = Q^T \geq 0\) and \(R = \varphi I_{F} > 0\) are weighting matrices associated to the magnitude of the state tracking error and control actions, respectively, while \(C\), composed of elements \(c_{r,s}(k)\), where

\[
c_{r,s}(k) = \begin{cases} 
1, & \text{if the density is tracked} \\
0, & \text{otherwise}
\end{cases}
\]

reflects the cells that are tracked. At first, we suppose to target only the cells at the bottleneck locations (e.g., in Figure 1, \(\rho_{3.1}, \rho_{3.2}\)).

The problem described by Equations 13, 11 is solved through a Linear Quadratic Regulator (LQR), under the assumptions that the original system is, at least, stabilisable and detectable (see Chapter 2 of (38)). As shown in (23), stabilisability is guaranteed for any network configuration, while, in order to guarantee detectability, it is necessary to control the density of each cell that does not have any other cell downstream. To account for this issue, we place an additional dummy cell immediately downstream of each lane-drop, imposing it, with an appropriate high penalty weight \(\alpha_{\hat{i}, \hat{j}}\), to have a density equal to zero. Note that, in the described case, the system is also observable. Further details are presented in (23).

The solution to the proposed LQR problem, obtained via Dynamic Programming in (23), results in the following feedback/feedforward control law

\[
u^*(k) = -K\hat{x}(k) + u_{ff},
\]

where

\[
K = (R + B^T PB)^{-1} B^T PA
\]

\[
P = C^T QC + A^T PA - A^T PB (R + B^T PB)^{-1} B^T PA
\]

\[
u_{ff} = K_{y} \hat{y} + K_{d} d
\]

\[
K_{y} = (R + B^T PB)^{-1} B^T (I - (A - BK)^T)^{-1} C^T Q
\]

\[
K_{d} = - (R + B^T PB)^{-1} B^T (I - (A - BK)^T)^{-1} P.
\]

Note that the optimal gain computed in Equation 16 and the Algebraic Riccati Equation (ARE) computed in Equation 17 are the same that can be found in classic Optimal Control books (see, e.g., (39)). Several methods have been proposed to compute efficiently the solution of the ARE (see, e.g., (39, 40)). Note also that, for practical implementation, we may allow for the (measured) inflow \(d\) to be time-varying, in which case the feedforward term \(u_{ff}\) in Equation 15 becomes also time-varying, obtaining (instead of Equations 15, 18)

\[
u^*(k) = -K\hat{x}(k) + u_{ff}(k)
\]

\[
u_{ff}(k) = K_{y} \hat{y} + K_{d} d(k).
\]
This corresponds to a model predictive control procedure, whereby the future inflow values are predicted to be equal to their current (measured) values.

It is important to highlight that the proposed feedback/feedforward control law is very effective for practical application since the computation of the feedback gain matrix \( K \) and of \( K_y \) and \( K_d \) is effectuated only once, offline; while online calculations are limited to few matrix-vector multiplications, as evidenced by Equations 21, 22.

A similar optimal regulation problem, without guarantee of regulation to an a priori prescribed set-point for state variables and non-zero mean disturbances, has been also considered in (41), where a different formulation for the feedforward term is obtained. In fact, our solution to the optimal control problem is obtained employing the Dynamic Programming principle, whereas (41) uses Lagrange multipliers. Although it is cumbersome to compare analytically the two control laws, they produce the same results in all the tested examples presented in this paper.

**Feedback control strategy for density distribution at bottlenecks**

We propose here an extended control strategy that, besides aiming at tracking the critical density (e.g., when demand is close to bottleneck capacity), also aims at distributing the vehicles at the bottleneck area, among the different lanes, according to a given policy.

To achieve this end, we modify the control law by choosing a time-varying set-point \( \hat{\gamma}(k) \) as a function of the network inflow: \( \hat{\gamma}(k) = \gamma(d(k)) \); where the function \( \gamma \) defines the pursued lane distribution policy. Thus, we maintain the feedback/feedforward control law in Equation 21, however, we replace the feedforward term of Equation 22 by

\[
  u_{ff}(k) = K_y \gamma(d(k)) + K_d d(k).
\]  

(23)

As an example, we show in Figure 2 possible functions for defining the set-points for the left \( \hat{\gamma}_L \) and right \( \hat{\gamma}_R \) lanes of a two-lane motorway. In this example, we impose that for low total inflow \( d_{tot} \) entering the motorway network, a higher amount of traffic is assigned to the left lane, by choosing \( \hat{\gamma}_L > \hat{\gamma}_R \) for \( 0 < d_{tot} < \bar{d}_{tot} \), where \( \bar{d}_{tot} \) is a flow value smaller than the bottleneck capacity \( d_{cap} \); while \( \hat{\gamma}_L = \rho_{L \text{cr}} \) and \( \hat{\gamma}_R = \rho_{R \text{cr}} \) for at \( d_{tot} = \bar{d}_{tot} \). As a result, we expect a higher outflow from the left lane when the incoming demand is lower than \( d_{tot} \), while both lanes should reach simultaneously their capacity (i.e., operating at their critical densities) when the overall demand approaches the bottleneck capacity. We would like to highlight that the proposed controller is capable to achieve a desired distribution of traffic based on any given functions, which would reflect different distribution policies. A constraint to be considered while defining such functions is that, in order to obtain the best traffic performance, the (per-lane) density set-points should be equal to the (per-lane) critical densities, when the inflow approaches to the bottleneck capacity.

As an alternative, the set-point \( \hat{\gamma}(k) \) may be varied via a total-density-dependent term \( \chi(\rho_{tot}(k))\rho_{tot}(k) \), where \( \chi \) is an opportunely defined function and \( \rho_{tot}(k) \) is the total (measured) density at the bottleneck area. In this case, \( \chi \) holds the portions of the total current density assigned to the corresponding lanes. Due to the involvement of \( \rho_{tot}(k) \), this leads factually to an additional (outer) feedback loop, which, however, has virtually no impact on the overall system stability, as numerical investigations have shown.

Finally, note that, all the proposed controllers are in the form of state-feedback regulators, which require availability of measurements for all state variables (densities for each cell) in real time. In the case of incomplete measurements, one may employ a traffic state estimator to produce
FIGURE 2 Possible functions $\psi(d(k))$ used to define in real-time the set-points $\hat{y}(k)$ at the bottleneck area as a function of the total inflow $d_{\text{tot}}(k)$ of the motorway network.

3 SIMULATION EXPERIMENTS
4 Nonlinear multi-lane traffic flow model
5 We proceed with performance evaluation of the proposed control strategies based on simulation experiments using a first-order traffic flow model based on (29). The model is used for reproducing the traffic behaviour for a multi-lane motorway and it features: (i) non-linear functions for the lateral flows of manually driven vehicles; (ii) a CTM-like (30) formulation for the longitudinal flows; and (iii) a non-linear formulation to account for the capacity drop phenomenon. We provide here a brief explanation of the employed model for self-completeness.

We consider the conservation law described in Equation 1. Lateral flows due to manual lane-changing are considered among adjacent lanes of the same segment, and corresponding rules are defined in order to properly assign and bound their values. The net lateral flows are computed as

$$ f_{i,j}(k) = l_{i,j,j+1}(k) - l_{i,j+1,j}(k), $$

where $l_{i,j,j}(k)$ is the lateral flow moving from cell $(i,j)$ to cell $(i,j)$ during time interval $(k,k+1]$ and $j = j \pm 1$; lateral flows $l_{i,j,j}(k)$ are computed according to

$$ l_{i,j,j}(k) = \min \left\{ 1, \frac{S_{i,j}(k)}{D_{i,j-1,j}(k) + D_{i,j+1,j}(k)} \right\} D_{i,j,j}(k) $$

$$ S_{i,j}(k) = \frac{L_i}{T} \left[ \rho_{i,j}^{\text{jam}} - \rho_{i,j}(k) \right] $$

$$ D_{i,j}(k) = \frac{L_i}{T} \rho_{i,j}(k) A_{i,j,j}(k) $$

$$ A_{i,j,j}(k) = \mu \max \left\{ 0, \frac{P_{i,j,j}(k) \rho_{i,j}(k) - \rho_{i,j}(k)}{P_{i,j,j}(k) \rho_{i,j}(k) + \rho_{i,j}(k)} \right\}, $$
Equation 25 accounts for the potentially limited space that may not be sufficient for accepting the lateral flow entering from both sides of a cell, where $S$ is the available space, in terms of flow acceptance, and $D$ is the lateral demand flow, which is computed via definition of the attractiveness rate $A$. The attractiveness rate is computed as a function of the densities for each pair of adjacent lanes; the factor $P$ affects the distribution of vehicles among lanes and should be calibrated to achieve the desired behaviour, e.g., using real data as in (45). Choosing a value $P = 1$ implies that drivers move always towards a faster lane (leading also to equal densities among lanes), but $P$ may also be tuned to reflect particular location-dependent effects where lateral flow may occur in the direction from a lower density to a higher one (e.g. upstream of on- and off-ramps, lane drop locations, etc.). Finally, parameter $\mu$ is a constant coefficient in the range $[0,1]$ reflecting the “aggressiveness” in lane-changing.

Longitudinal flows are the flows generated in a segment and moving to the next downstream one, while remaining in the same lane. We employ a Godunov-discretised scheme similar to the one proposed in (29), using however the non-linear exponential function proposed in (46) to obtain a more realistic behaviour at undercritical densities. The model accounts also for the capacity drop phenomenon via a linearly decreasing demand function for over-critical densities; in addition, other modelling approaches can be employed to improve the capability of reproducing capacity drop, obtaining comparable results (see, e.g., (47, 48)). More details and calibration results related to this model are presented in (29, 45). Formally, the complete formulation for longitudinal flows reads

$$q_{i,j}(k) = \min \left\{ Q_{i,j}^D(k), Q_{i+1,j}^S(k) - d_{i,j}(k) \right\}, \quad (29)$$

where

$$Q_{i,j}^D(k) = \begin{cases} v_{\text{max}}^{i,j} \exp \left[ -\frac{1}{\alpha} \left( \frac{\rho_{i,j}(k)}{\rho_{i,j}^c} \right)^\alpha \right] \rho_{i,j}(k), & \text{if } \rho_{i,j}(k) < \rho_{i,j}^c \\ \rho_{i,j}^c \left( \rho_{i,j}(k) - \rho_{i,j}^{\text{jam}} \right) + \gamma Q_{i,j}^{\text{cap}}, & \text{otherwise} \end{cases} \quad (30)$$

$$Q_{i+1,j}^S(k) = \begin{cases} Q_{i+1,j}^{\text{cap}}, & \text{if } \rho_{i+1,j}(k) < \rho_{i+1,j}^c \\ w_{i+1} \left( \rho_{i+1,j}^{\text{jam}} - \rho_{i+1,j}(k) \right), & \text{otherwise}. \end{cases} \quad (31)$$

Parameter $v_{\text{max}}^{i,j}$ denotes the free speed, $Q_{\text{cap}}$ is the capacity flow, $\rho^c$ is the critical density (i.e., the density at which the capacity flow occurs), $\gamma$ is a capacity drop coefficient within $[0,1]$, while $\alpha = \left( \ln \frac{Q_{\text{cap}}}{v_{\text{max}}^{i,j} P} \right)^{-1}$.

**Network description and the no-control case**

We consider a hypothetical motorway stretch to test and evaluate the performance of the proposed strategy. In particular, we consider the network depicted in Figure 3, which is composed of 7 segments; segments 1, . . . , 5 feature three lanes, while segments 6 and 7 feature only two lanes, with a lane-drop located downstream of cell (5, 1). All segments are characterised by the same length $L_i = 0.5 \text{ km}$, while we define a simulation step $T = 10 \text{ s}$. Different lanes feature different parameters, specifically a different Fundamental Diagram, which may reflect different traffic composition (e.g., a higher rate of heavy vehicles reducing the capacity of a specific lane); the used values are shown in Table 1.
FIGURE 3 The motorway stretch used for testing and evaluating the proposed control strategy.

### TABLE 1 Parameters used in the nonlinear multi-lane traffic flow model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Lanes $j = 1, 2$</th>
<th>Lane $j = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_{\text{max}}$ [km/h]</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{\text{cap}}$ [veh/h]</td>
<td>1800</td>
<td>2400</td>
</tr>
<tr>
<td>$\rho^c$ [veh/km]</td>
<td>32</td>
<td>36</td>
</tr>
<tr>
<td>$\rho_{\text{jam}}$ [veh/km]</td>
<td>120</td>
<td>160</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.65</td>
<td>0.65</td>
</tr>
<tr>
<td>$P$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Traffic demand profiles are defined for a simulation horizon $K = 480$ (80 min), as shown in Figure 4. Note that the overall demand entering the network is, at its peak, roughly equivalent to the total capacity of segment 5, i.e., the bottleneck capacity.

Running the macroscopic model described by Equations 1, 24–31 without the use of any control actions produces eventually traffic congestion starting at the lane-drop area, due to non-optimal spontaneous lane-changes of vehicles. Inspecting the contour plots shown in Figure 5 (top), we can see that the density increases first in lane 1 (the one that is dropping) at around $t = 20$ min due to the high demand arriving in the lane-drop area, while vehicles try to merge first into lane 2, and, due to the fact that density increases also in this lane, eventually also into lane 3. In particular, most lane-changes take place within segments 4 and 5, while a small amount of lane-changes take place within segment 6 and there are virtually no lane-changes in the upstream segments (see Figure 6 (top)). We recall that, according to Equation 28, with $P_{i,j,j} = 1$, the lane-changing model acts towards the homogenisation of the densities between adjacent lanes. The detrimental effects of the congestion worsen as a consequence of the occurring capacity drop, which is here triggered by overcritical densities at both lanes of segment 5, causing a reduction of the outflow in both lanes during the high-demand period, as shown in Figure 7 (top).

The created congestion spills back covering all lanes of segments 4 and 5 (see Figure 5). As numerical evaluation criterion we employ the Total Travel Time (TTT) over a finite time horizon $K$, defined, as in (27), as

$$\text{TTT} = T \sum_{k=0}^{K} \sum_{i=0}^{N} \sum_{j=m_i}^{M_i-1} L_i \rho_{i,j}(k),$$

obtaining, for the presented no-control case, a resulting overall $\text{TTT} = 186.7 \text{ veh} \cdot \text{h}$. 

TRB 2017 Annual Meeting Paper revised from original submittal.
FIGURE 4 The traffic demand $d$ entering segment 1 (see Figure 3).

Application of the control strategy with constant set-points

We proceed now to the evaluation of the optimal control strategy with constant set-points using the previously described motorway scenario. We define as “application area”, namely the portion of network where we apply our designed strategy, the area from segment 3 to segment 6 (see Figure 3). We use the outflow of the segments immediately upstream of the application area $q_{2,j}$ as demand $d$. A dummy cell $(6,1)$ is added immediately downstream of the lane-drop in order to ensure system observability. The set-point considered in the LQR includes thus the three cells in segment 6.

According to the network topology and setting a constant speed $\bar{v} = 90 \text{ km/h}$ and cost weights $Q_{i,j} = 1$, for $i = j = 2, 3$; $Q_{i,j} = 100$, for $i = j = 1$; $Q_{i,j} = 0; \forall i \neq j$; $\phi = 10^{-5}$ (obtained after some manual tuning of the controller aiming at achieving an efficient and smooth response), we compute (offline) the gains according to Equations 16, 17, 19, 20.

Assuming the critical densities at the controlled area to be known, we build the set-point vector $\hat{y}$ to consist of $\hat{r}_{6,2} = 32 \text{ veh/km}$, $\hat{r}_{6,3} = 36 \text{ veh/km}$, while for the additional dummy segment we define $\hat{r}_{6,1} = 0 \text{ veh/km}$.

Lateral flows $f_{i,j}$ are computed as $u^*$, via the control law (Equation 21), and are then applied directly in the conservation law (Equation 1) of the simulation model, while longitudinal flows $q_{i,j}$ are obtained from Equations 29–31 as in the no-control case.

From inspection of the resulting contour plots in Figure 5 (middle), we can see that the controller is capable of avoiding the creation of congestion. This is due to the fact that, during the period characterised by high demand, the density at the bottleneck area is maintained at its critical value. The optimal lateral flows are distributed quite homogeneously within the whole application area (see Figure 6 (middle)), thus avoiding high lane-changing flows close to the lane-drop location. Moreover, since all densities remain undercritical, the capacity drop phenomenon
FIGURE 5  Contour plots of densities in the no-control case (top), when the control strategy with constant set-points is applied (middle), and when the proposed feedback control strategy for density distribution is applied (bottom).

is not appearing, and the system operates at the bottleneck capacity during the whole peak period (see Figure 7 (middle)). Within this scenario, we obtain a $\text{TTT} = 145.7 \text{ veh} \cdot \text{h}$, which is a 22% improvement with respect to the no-control case.

However, as we can see from Figure 8 (left), at the bottleneck area, the flow exiting lane 3 is always higher than the flow exiting lane 2, for any value of total flow. This is due to the higher value of critical density used as constant set-point within the application of this control strategy.

Application of the proposed feedback control strategy for density distribution at bottlenecks

We now test the proposed control strategy aiming at distributing the total density at a bottleneck area, among the different lanes, according to a given policy. The set-point vector $\hat{y}(k)$ is computed via the functions depicted in Figure 2, employing a quadratic form for $\hat{\rho}_{6,2}(k)$ and a linear term for
FIGURE 6  Contour plots of net lateral flows in the no-control case (top), when the control strategy with constant set-points is applied (middle), and when the proposed feedback control strategy for density distribution is applied (bottom).
FIGURE 7 The flow exiting from lanes 2 (left) and 3 (right) of segment 5 (blue lines) and the corresponding capacity flow (red lines). In the no-control case (top), the capacity drop mechanism is triggered and the outflow drops from the capacity flow; whereas, when the control strategy with constant set-points (middle) or the proposed feedback control strategy for density distribution (bottom) are applied, the capacity drop phenomenon is avoided and the outflow, during the peak period, is close to the bottleneck capacity.
\( \hat{\rho}_{6,3}(k) \) according to

\[
\hat{\rho}_{6,2}(k) = \begin{cases} 
\frac{1}{v_{\text{tot}}} [d_{\text{tot}}(k)]^2 + \frac{\sqrt{\rho_{6,2}^{\text{cr}}}}{v_{\text{tot}}} d_{\text{tot}}(k), & \text{if } d_{\text{tot}}(k) \leq \tilde{d}_{\text{tot}} \\
\rho_{6,2}^{\text{cr}}, & \text{otherwise,}
\end{cases}
\]

\[
\hat{\rho}_{6,3}(k) = \begin{cases} 
\frac{\sqrt{\rho_{6,3}^{\text{cr}}}}{d_{\text{tot}}(k)}, & \text{if } d_{\text{tot}}(k) \leq \tilde{d}_{\text{tot}} \\
\rho_{6,3}^{\text{cr}}, & \text{otherwise,}
\end{cases}
\]

where

\( d_{\text{tot}} = \frac{4}{5} d_{\text{cap}}. \)

We maintain the same configuration for the controlled system as in the previous case, computing the lateral flow as \( u^l \) via the feedback/feedforward control law in Equation 21, however using Equation 23 for computing the feedforward term.

Similarly to the previous case, the resulting contour plots in Figure 5 (bottom) illustrate that the controller also avoids congestion and hence the capacity drop phenomenon during the whole peak period (see Figure 7 (bottom)), while lateral flows are distributed quite homogeneously within the whole application (see Figure 6 (bottom)). For this scenario, we obtain a \( TTT = 146.7 \text{ veh} \cdot \text{h}, \) which is a 21.4\% improvement with respect to the no-control case.

In this case, however, we can see from Figure 8 (right) that, at the bottleneck area, the flow exiting lane 2 is higher than the flow exiting lane 3 for lower values of total flow (i.e., when the total flow is lower than about 3500 veh/h); whereas, for higher values of total flow, the flow in lane 3 exceeds the flow in lane 2 until capacity flow is reached simultaneously. Note that in Figure 8 (right) there are three equilibrium values (circled) for outflows at each lane, which can be identified as areas where the marks appear thicker, which are representative of the respective periods of simulation characterised by low, intermediate, and high traffic demand (see Figure 4). The observed behaviour is in full accordance with the goals of the employed policy for lane distribution.
CONCLUSIONS

In this paper we presented an extended version of an optimal control strategy for lane-changing-based traffic control at bottleneck locations, which we previously proposed in (23), by including, together with the capability to operate a motorway traffic system at its capacity, the possibility to distribute the traffic at the bottleneck area, among the different lanes, according to a given policy. Simulation results demonstrate the effectiveness of the proposed control strategy in improving traffic performance, while also pursuing a prescribed lane flow distribution at the bottleneck area.

We are currently extending this methodology to account for unmeasured demand flows and incomplete measurements, as well as to incorporate a mainstream or ramp flow control strategy. Moreover, we are looking into the case of mixed traffic, where manual vehicles may not receive or may not follow the prescribed lane-changing commands.

ACKNOWLEDGEMENT

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 321132, project TRAMAN21.
REFERENCES


