Transportation Research Part B 45 (2011) 769–781

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Increasing the capacity of signalized intersections with separate left turn phases

Yiguang Xuan, Carlos F. Daganzo *, Michael J. Cassidy

Institute of Transportation Studies, University of California, Berkeley, CA 94720, United States

ARTICLE INFO

Article history:
Received 8 July 2010
Received in revised form 7 February 2011
Accepted 8 February 2011

Keywords:
Traffic signal capacity
Pre-signals
Left turn phases

ABSTRACT

A separate turn phase is often used on the approach leg to an intersections with heavy left turns. This wastes capacity on the approach because some of its lanes cannot discharge during its green phases. The paper shows that the problem can be eliminated by reorganizing traffic on all the lanes upstream of an intersection using a mid-block pre-signal. If drivers behave deterministically, the capacity that can be achieved is the same as if there were no left turns. However, if the reorganization is too drastic, it may be counterintuitive to drivers. This can be remedied by reorganizing traffic on just some of the available lanes. It is shown that such partial reorganization still increases capacity significantly, even if drivers behave randomly and only one lane is reorganized. The paper shows how to optimize the design of a pre-signal system for a generic intersection. It also identifies both, the potential benefits of the proposed system for a broad class of intersections, and the domain of application where the benefits are most significant.

1. Introduction

Left turns can be a significant hindrance to the smooth flow of traffic in networks involving at-grade intersections. If left turns are sufficient in number, separate left turn phases are typically introduced at signalized intersections to handle the flow (Newell, 1989). The problem with left turns is also indirectly addressed by generic optimization methods devised to determine the duration of signal phases (e.g., Webster, 1958; Allsop, 1972), which were later generalized to include the grouping of streams into phases (e.g., Improta and Cantarella, 1984; Gallivan and Heydecker, 1988; Silcock, 1997), and more recently to determine the number of lanes that should be made available to the streams (Wong and Wong, 2003).

The above-cited works are concerned with the reduction of vehicle delays when the intersection approach is undersaturated; i.e., when it has sufficient capacity to serve the demand. Vehicle-actuated (VA) signal control is often used to further reduce this delay by dynamically allocating green times to accommodate cyclic fluctuations in demand. Although VA control can reduce delay on undersaturated approaches (Greenough and Kelman, 1999; Peck et al., 1999; Sussman et al., 2000), it is well known that VA signals can do little to increase the capacity of oversaturated approaches (Newell, 1989; Lo and Chow, 2004). The effects of oversaturation can only be eliminated by reducing demand or increasing capacity. If nothing is done, oversaturation can cause gridlock at the network level (Daganzo, 2007).

Left turns in large numbers contribute to the oversaturation, because they require separate green phase allocations and these sub-phases reduce intersection capacity.[1] These capacity problems are often avoided in practice by banning and re-routing the offending left-turns. This can be done in different ways; e.g., with median U-turns, jughandles, superstreets, paired intersections, quadrant roadways, and bowties (see e.g. ATTAP, 2006; Rodegerdts et al., 2004). These strategies are good

* Corresponding author.
E-mail address: daganzo@ce.berkeley.edu (C.F. Daganzo).

[1] Clearly, only part of the intersection cross section is used to discharge vehicles during each of these green sub-phases.

0191-2615/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.trb.2011.02.009

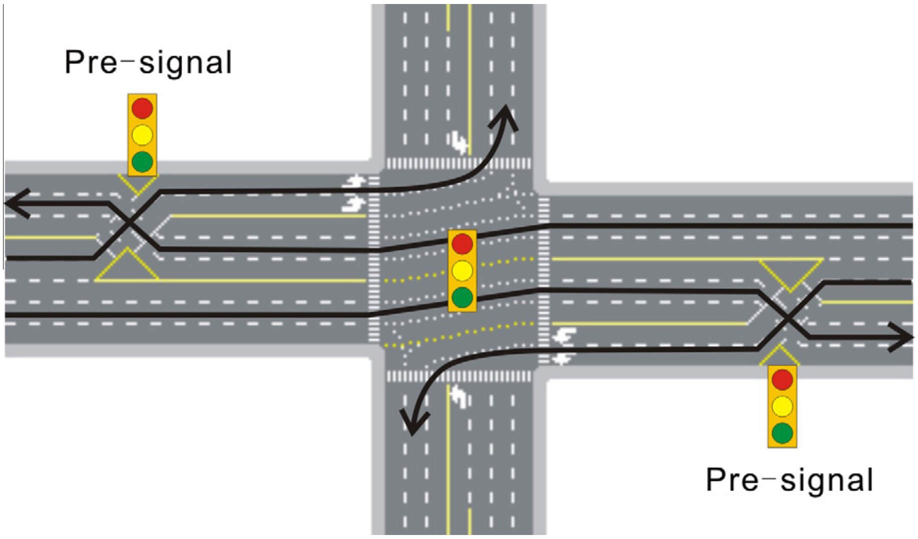

Fig. 1. Continuous flow intersection. Arrows show the movements allowed during the intersection's green phase for the horizontal directions.

because they eliminate both, the left turns and the need for left turn phases, thereby increasing the intersection capacity. But the strategies also force left-turning vehicles to go through the intersection multiple times, increasing demand there. This offsets some of the benefit. In addition, most re-routing designs involve construction and require substantial space, which may not be available.

In view of this, this paper proposes a way of increasing capacity without banning left turns or modifying the intersection in a way that requires major construction. The focus of the analysis is a generic approach leg to a 4-way intersection controlled by a pre-timed traffic signal. The approach receives two green sub-phases: one for left turns only, and the other for the exclusive use of through and right-turn movements.

Currently, intersections of this type do not operate efficiently because during the protected left-turn sub-phase, only the lanes devoted to left turns discharge. Conversely, during the through sub-phase only the through (and right turn) lanes discharge. This is inefficient because many lanes go unused during the green phases. The problem is compounded when lanes are shared by through and left-turn vehicles, which cannot fully discharge during either sub-phase due to blocking. Imagine how much more capacity could be gained if all lanes could be made to fully discharge during both sub-phases. The paper will show how this objective can be realized by dynamically re-organizing traffic upstream of the intersection with a pre-signal.

The idea of pre-signals is not new. They have been used in Europe to give buses priority for a long time (Oakes et al., 1994). Pre-signal systems for both, buses and bicycles have been studied in Wu and Hounsell (1998) and Xuan et al. (2009), respectively. Pre-signals have also been used in connection with left turns; see Fig. 1, which illustrates the "continuous flow intersection (CFI)" concept (Al-Salman and Salter, 1974; Goldblatt et al., 1994). This concept resolves the left-turn vs. opposing-through conflict at the pre-signals, as shown in the figure. Since the left-turn vehicles and opposing-through vehicles are reorganized to avoid conflicts at the intersection, they can be served in a single green phase, and this increases capacity. Interesting as it is, the CFI concept still has some drawbacks, however: (a) it requires permanent changes in layout, which cannot be turned on and off as needed; (b) it wastes lane resources unless the demands for left and through vehicles happen to be in the same balance as the lanes are marked; (c) it cannot be surgically deployed on a single approach of a 4-way intersection—at least 2 approaches need to be involved; and (d) it precludes bus stops from being located on the far side of the intersection.

The proposed design is similar in spirit but overcomes these problems. Since it can be deployed on a single approach leg, a single approach will be the object of analysis. Section 2 describes the concept; Section 3 the optimum signal settings and capacity formulae; and Section 4 discusses the results, including the spatial requirements of the system.

2. The concept

The basic idea is depicted in Fig. 2a. Flow is left to right. Note the pre-signal and its (dark) stop line. Upstream of the pre-signal, lanes are marked by movement. These markings segregate left-turning vehicles (LVs) and through/right-turning vehicles (TVs) onto separate sets of lanes.[2] The figure does not show right turns for simplicity of exposition; but the system performs just the same if they are included.

[2] We assume that drivers can maneuver into their desired lanes as they approach the pre-signal, just as we would ordinarily assume that drivers can maneuver into their desired lanes as they approach a conventional signalized intersection.

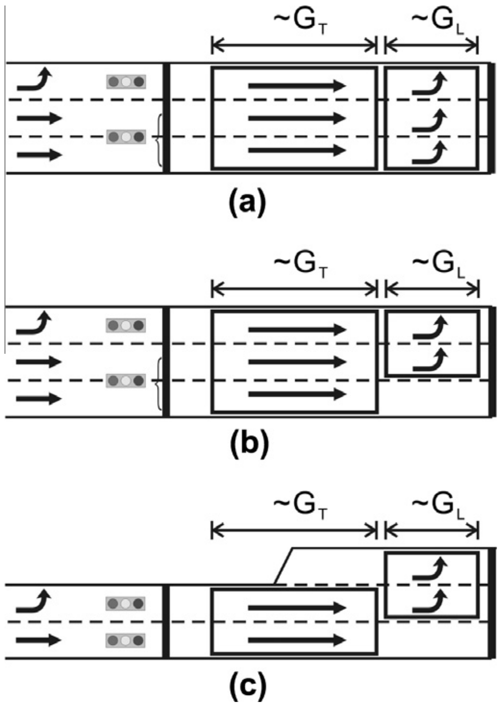

Fig. 2. The tandem intersection concept: (a) full tandem; (b) partial tandem; (c) general case. Rectangles indicate the size of the LV and TV queues in each lane, and their order; they do not always indicate the actual extent of these queues.

The pre-signal runs on the same cycle as the intersection signal and alternates giving green time to the two sets of lanes. Downstream of the pre-signal, variable message signs (VMSs) instruct drivers of both vehicle types which lanes to use – possibly all of them as shown in Fig. 2a. It may be desirable to change these instructions as traffic conditions change. Therefore, VMSs are preferred to lane markings because the former can be changed with time. The area between the pre-signal and the signal will be called the "sorting area". This area is intended to contain those transient queues created by the different offsets and phases of the signal and the pre-signal. The sorting area should be long enough to ensure that these queues do not spill back to the pre-signal.

Here is how the system of Fig. 2a would operate for one cycle if at the intersection the LV-sub-phase leads the TV-sub-phase. The pre-signal starts its cycle by giving the green to the LVs while the intersection signal is red. These LVs advance into the sorting area using all the lanes and wait at the intersection stop line. The pre-signal then turns green for the TVs, which line up behind the LVs, again using all lanes.[3] With this arrangement, vehicles are efficiently positioned to discharge using all the lanes when the intersection green phases start: first the LV-sub-phase, then the TV-sub-phase. Of course, if the LV-sub-phase lags the TV-sub-phase then the order is reversed and the pre-signal starts its cycle with the TVs.

The duration of the pre-signal phases (for LVs and TVs) should be chosen so that the number of vehicles of either type discharged by the pre-signal at saturation is close to, but does not exceed, the maximum number that all the intersection lanes can jointly discharge during the corresponding intersection sub-phase. In this way queues of either vehicle type clear at the signal in every sub-phase so that blocking is avoided when sub-phases change. The pre-signal phases should also ensure that if there is sufficient demand, then the discharged numbers should be close to this maximum. In this way both signal sub-phases will be nearly saturated. This idea is illustrated in Fig. 2a by the two rectangular blocks in the sorting area. These blocks represent the space needed for the maximum number of vehicles that the signal can discharge in its two sub-phases. The sub-phase durations shall be denoted G_L and G_T, where the single-letter subscript refers to either LTs or TVs. Lengthwise, these blocks span the maximum number of vehicles that can discharge from a single lane in a sub-phase. Therefore, their lengths are proportional to G_L and G_T, as shown.

An advantage of the proposed pre-signal method when compared with approaches that require fixed infrastructure and/or markings is that the phases and sub-phases of the two signals can be changed with the time-of-day to be in balance with the time-varying percentage of turns. This produces (nearly) the same capacity as if left turns had been banned. A disadvantage is that drivers are sometimes asked to be in a counterintuitive lane; e.g., note from Fig. 2a that some LVs occupy the intersection's right lane.

[3] We assume that drivers will distribute themselves evenly over the lanes available to them.

To alleviate this drawback, the full-tandem concept of Fig. 2a can be moderated with ''partial-tandem'' designs in which only some of the lanes in the sorting area are shared by both types of vehicles; see Fig. 2b, where only two lanes are shared in tandem. To avoid confusion with other types of shared lanes, these lanes will be called ''tandem lanes'' from now on. The disadvantage of the partial-tandem concept is that it does not engage all the lanes fully in each sub-phase. To balance user-friendliness with capacity needs, one can define for any application the smallest possible set of tandem lanes that will meet demand.

Since the full-tandem case is a special case of partial tandem, we will from now on focus on the latter, often dropping the word ''partial''. The most general case (illustrated in Fig. 2c) does not even require that the number of lanes at the signal and the pre-signal be equal.

3. Capacity analysis

The capacity of tandem and conventional pre-timed intersection designs are compared here.[4] Capacity formulae and capacity-maximizing timing plans are developed, first for an idealized best case scenario in which vehicles behave deterministically, and then for a more realistic scenario that captures stochastic phenomena. Although the corrections introduced for the latter reduce the ideal capacity, results show that the tandem design typically outperforms the conventional design nonetheless. The outperformance is considerable when left turns are numerous.

In what follows the letters g, n, q shall denote the green times, number of lanes, and vehicle discharge flows under various scenarios and at different locations. The superscript ''0'' will indicate a conventional design without a pre-signal, and no superscript the tandem design. Subscripts ''L'' and ''T'' will indicate the vehicle type, and no subscript the total. Upper and lower cases for g and n will refer to the intersection and the pre-signal respectively. Note that $N_L + N_T - N$ is the number of tandem lanes, which can range from 0 to N; and that $n = n_L + n_T$, since there are no shared lanes upstream of the pre-signal.

To simplify the expressions, we shall use the cycle length as the unit for time, and the average number of TVs (or LVs) that one lane can discharge in one cycle's worth of green time as the unit of TV (or LV) quantity. In this system of units, the green phases g_L, g_T, G_L, $G_T < 1$ are also the fractions of green, and the saturation flows per lane of both vehicle types are 1. Figs. 2b–c use rectangles just as Fig. 2a to denote the maximum number of vehicles that the signal can discharge in its two sub-phases. These rectangles would have the same lengths in all the figures if the sub-phase durations were the same in the three cases, but these rectangles span fewer lanes in Figs. 2b and c. This reduced rectangular width illustrates why the partial tandem system can sustain smaller flows.

Given as parameters are: the total green time available to the approach (G); the fraction of left-turns to be accommodated ($l = q_L/q$); the original approach geometry, with no shared lanes anywhere ($N_L^0, N_T^0, N^0 = N_L^0 + N_T^0, n_L, n_T, n = n_L + n_T$); and the modified lane definitions for the tandem design in the sorting area ($N = N^0, N_L, N_T,$ with $N \leqslant N_L + N_T \leqslant 2N$).

For both the conventional and tandem designs, we first solve the ''optimal timing problem''; i.e., we find the maximum total flow that the approach can sustain for a given G by optimally choosing the intersection sub-phase durations (G_L^0, G_T^0, G_L, and G_T), and for the tandem design the pre-signal phase durations (g_L and g_T) too. In doing so, we shall assume that the lost time at the pre-signal can be neglected; i.e., that the pre-signal green phases can span the whole cycle: $g_L + g_T \leqslant 1$. This seems reasonable as a first approximation since the conflicts resolved at the pre-signal involve vehicles traveling in the same direction. Likewise, we shall also assume that G is an ''effective green'' so we can choose any sub-phases such that $G_L^0 + G_T^0 \leqslant G$ and $G_L + G_T \leqslant G$. We shall ignore the impact of this choice on the opposing flows, which is reasonable if these flows are low. Otherwise, a couple of capacity constraints should be added to the formulation about to be presented. Though these additions are straightforward, the effect of high opposing flows will not be evaluated in this paper in the interest of brevity.

With the optimum timing problem as a building block, we then solve the ''optimal design problem''. For this problem, the lane assignments are no longer treated as parameters, but as decision variables.

3.1. Deterministic analysis

It is assumed in this section that drivers discharge from the traffic signals with deterministic headways. The timing problem is presented first.

3.1.1. Optimal timing

The maximum flow of the conventional design is given by the following simple linear program, which we denote as ''CLP'' (conventional linear program). The geometry (i.e., the number of lanes) is given, and the green sub-phase durations are the decision variables.

$$(\text{CLP}) \max_{G_L^0, G_T^0} q^0$$

$$q_L^0 = q^0 l; \quad q_T^0 = q^0(1-l) \tag{1.1}$$

[4] Recall the VA designs have similar capacities to the pre-timed designs.

$$G_L^0 + G_T^0 \leqslant G \tag{1.2}$$

$$q_L^0 = G_L^0 N_L^0 \leqslant n_L \tag{1.3}$$

$$q_T^0 = G_T^0 N_T^0 \leqslant n_T \tag{1.4}$$

$$G_L^0, G_T^0 \geqslant 0 \tag{1.5}$$

Eq. (1.1) ensures that the intersection serves the correct balance of both vehicle classes; (1.2) ensures that only the available green time is used; and (1.3) and (1.4) that both the signals and the upstream approach lanes can sustain the flows. The inequalities in (1.3) and (1.4) are redundant if $n_L \geqslant GN_L^0$ and $n_T \geqslant GN_T^0$. This is almost always the case.

Consideration shows that if n_L and n_T do not enter the picture then the optimum solution of the CLP is:

$$q^0 = \frac{G}{l/N_L^0 + (1-l)/N_T^0} \tag{2.1}$$

$$q_L^0 = q^0 l \tag{2.2}$$

$$q_T^0 = q^0(1-l) \tag{2.3}$$

$$G_L^0 = q^0 l/N_L^0 \tag{2.4}$$

$$G_T^0 = q^0(1-l)/N_T^0 \tag{2.5}$$

Eq. (2.1) gives the capacity and (2.4) and (2.5) the signal sub-phase durations. As expected, (1.2) is binding.

The maximum flow of the tandem design can also be found with a similar LP, which shall be called "TLP". The main two differences are that now there are four green phase durations to be determined, and that the flows of the two vehicle types are now constrained both by the signal and the pre-signal. If the storage area is long enough to prevent any transient queues at the signal from backing up to the pre-signal, their capacities can be evaluated independently. This assumption is made in the LP below:

$$(\text{TLP}) \max_{G_L, G_T, g_L, g_T} q$$

$$q_L = ql; \quad q_T = q(1-l) \tag{3.1}$$

$$G_L + G_T \leqslant G \tag{3.2}$$

$$g_L + g_T \leqslant 1 \tag{3.3}$$

$$q_L = G_L N_L = g_L n_L \tag{3.4}$$

$$q_T = G_T N_T = g_T n_T \tag{3.5}$$

$$G_L, G_T, g_L, g_T \geqslant 0 \tag{3.6}$$

New in this program are constraint (3.3) and the last members of constraints (3.4) and (3.5). Although the TLP appears to be more constrained due to these additions, this is not necessarily so because the tandem design always has more lanes for at least one vehicle type; i.e., $N_L \geqslant N_L^0, N_T \geqslant N_T^0$, and $N_L + N_T > N_L^0 + N_T^0$.

As before, the reader can verify that the solution of the TLP is:

$$q = \min\left\{\frac{G}{l/N_L + (1-l)/N_T}, \frac{1}{l/n_L + (1-l)/n_T}\right\} \tag{4.1}$$

$$q_L = ql \tag{4.2}$$

$$q_T = q(1-l) \tag{4.3}$$

$$G_L = ql/N_L \tag{4.4}$$

$$G_T = q(1-l)/N_T \tag{4.5}$$

$$g_L = ql/n_L \tag{4.6}$$

$$g_T = q(1-l)/n_T \quad (4.7)$$

The second term of (4.1) is the maximum flow that can pass through the pre-signal, and the first term the maximum flow through the signal. For an intersection without turn pockets $\left(N_L^0 = n_L \text{ and } N_T^0 = n_T\right)$, the second term is greater than (2.1) by a factor of $1/G$. Furthermore, since $N_L \geqslant N_L^0, N_T \geqslant N_T^0$, and $N_L + N_T > N_L^0 + N_T^0$, the first term of (4.1) is also greater than (2.1); perhaps considerably so. Thus, for any intersection without turn pockets the tandem system improves capacity.

The improvements for typical cases can be quite large. For example, assume that: $G = 1/2; l = 1/3; N_L^0 = 1; N_T^0 = 2$; and $n_L = 1$, $n_T = 2$, $N_L = N_T = 3$ as in Fig. 2a. Then, for the conventional design, (2.1), (2.2), (2.3) yield: $q_L^0 = 1/4; q_T^0 = 1/2$; and $q^0 = 3/4$; and for the tandem design (4.1), (4.2), (4.3) yield: $q_L = 1/2$; $q_T = 1$; and $q = 3/2$. Thus, capacity is increased by 100%. The same calculations for the partial design of Fig. 2b and c show capacity increases of 71% and 33% respectively. These results are typical. Although it is possible to construct examples with turn pockets where the tandem design would reduce capacity, these examples always involve unreasonably large values of G and/or small values of n/N.

3.1.2. Optimal design

The above assumes that the lane assignments are given and only the durations of the pre-signal phases and signal sub-phases can be adjusted. In reality lane assignments can be changed in response to changing demand. It is therefore of some interest to know the capacity that can be achieved when both the green-phase durations and the lane designations ($n_L, n_T, N_L, N_T, N_L^0$, and N_T^0) are optimally set. This can be done both for the conventional and tandem configurations by adding the following constraints to the CLP and TLP:

$$n_L + n_T = n \quad \text{for CLP and TLP} \quad (5.1)$$

$$N_L^0 + N_T^0 = N \quad \text{for CLP} \quad (5.2)$$

$$N_L, N_T \leqslant N \quad \text{for TLP} \quad (5.3)$$

$$N_L + N_T - N = N_{LT} \quad \text{for TLP with } N_{LT} \text{ tandem lanes} \quad (5.4)$$

With the new variables, some of the constraints in the CLP and the TLP become non-linear. Furthermore, since the new variables are natural numbers, the resulting problems are mixed-integer non-linear programs. In practical cases the number of feasible lane designations arising from (5) is small. Therefore, these non-linear programs can be easily solved by evaluating (2.1) or (4.1) for all feasible lane designations and choosing those that maximize the result.

To identify the proper application domain for the tandem method, a battery of these generalized CLP and TLP problems were solved. Four basic approach configurations were examined: approaches with 2 and 3 lanes, with and without 1-lane turn pockets. Their capacities were evaluated for all possible values of the input parameters (G and l), both under conventional and tandem control. For the sake of brevity, the latter was only analyzed for the case in which only one tandem lane is used; i.e., where its potential benefit is least. Fig. 3 shows the result. Contour lines show the ratio of the partial tandem capacity over the maximal possible flow (GN). Shading shows the ratio of the partial tandem capacity over the conventional capacity, with lighter shades denoting greater benefits reaped by the tandem design. These figures can also be used to determine if a given pair of flows can be served with a given G. To do so, go to the appropriate figure and determine from the point with the relevant G and l whether the given total flow is larger or smaller than the capacity displayed for the point's contour.

Note that when G is very large (approaching 1) the tandem system performs worse than the conventional. This happens because the pre-signal then becomes the binding constraint. But for common settings with G around 0.5, the capacity increase is about 20–30%. This is encouraging because the benefit is obtained with only one tandem lane. On the other hand, the results in this sub-section are optimistic because they assume that vehicles discharge from the intersection at a deterministic rate. The next sub-section obtains more realistic results by relaxing this assumption.

3.2. Stochastic considerations

The concepts discussed below are easier to understand if the cycle time (C) and the average discharge headway (H) are explicitly used as variables. Therefore, the special system of units we have been using, where these two variables took the value "1", will not be used. Instead the results will be derived for an arbitrary system and then converted to the special system at the end.

Refer to Fig. 2a (full tandem) and consider what would happen if the LVs in one of the sorting area's lanes were unable to fully discharge in their signal sub-phase (call this a lane failure) due to the randomness in discharge headways. Assume now that vehicles do not change lanes once sorted and consider what happens. The residual LVs in the failed lane would have to stop and wait for the next sub-phase. No TVs in that lane would discharge in their sub-phase. Instead, these TVs would have to wait until the residual LVs have cleared the intersection.[5] Note that the next cycle will discharge these queues and nobody else. Thus a full cycle would have been wasted for both vehicle types.

[5] In reality, some of these TVs may change lanes to take advantage of unused capacity in neighboring lanes. Thus, our assumption of no lane changing is conservative.

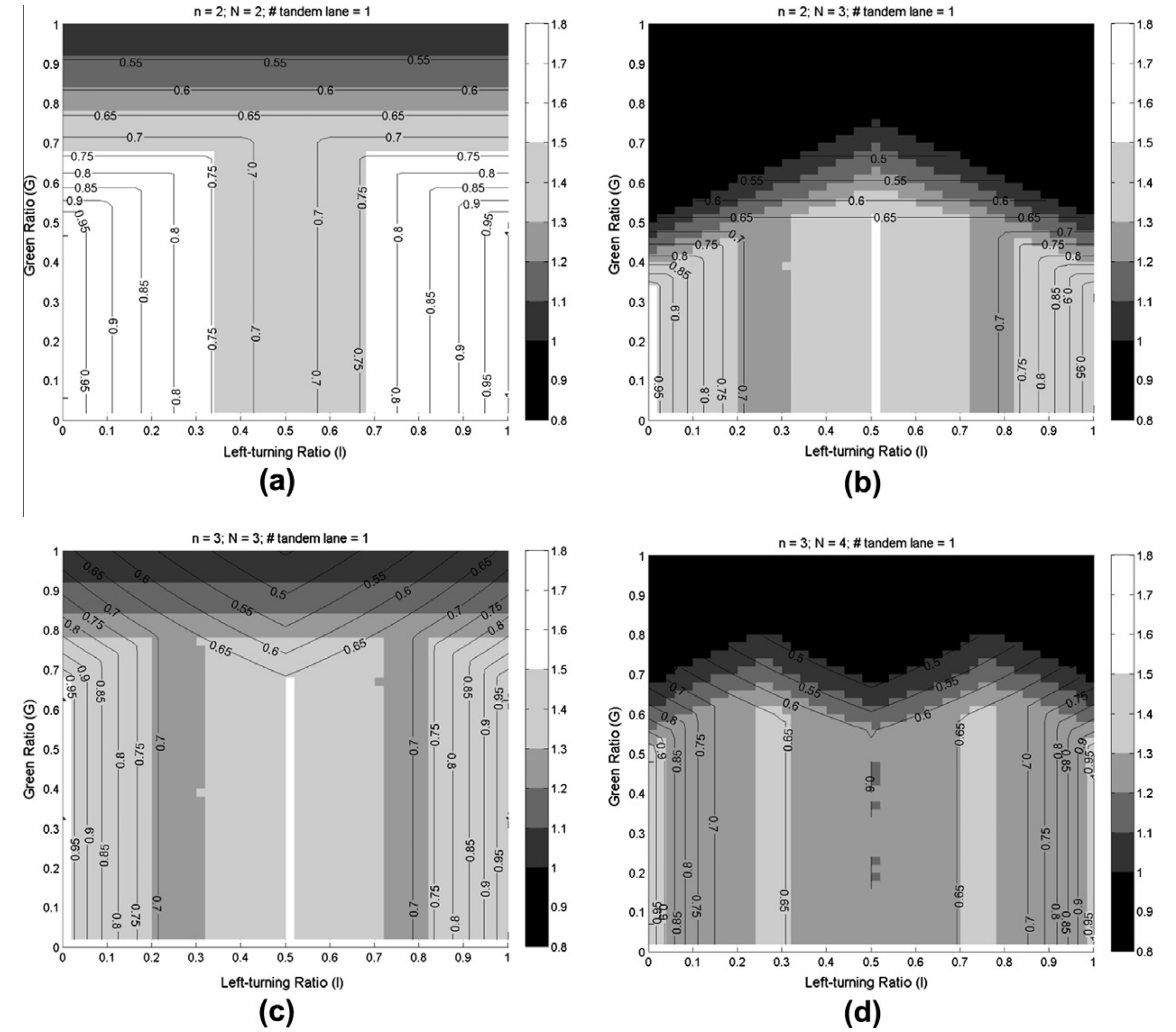

Fig. 3. Comparison of the deterministic version of capacities of conventional and tandem design. Contour lines show the ratio of the tandem capacity over the maximal possible flow; shading shows the ratio of the tandem capacity over the conventional capacity: (a) 2 lanes with no turn pocket ($n = N = 2$), 1 tandem lane; (b) 2 lanes with 1-lane turn pocket ($n = 2$, $N = 3$), 1 tandem lane; (c) 3 lanes with no turn pocket ($n = N = 3$), 1 tandem lane; (d) 3 lanes with 1-lane turn pocket ($n = 3$, $N = 4$), 1 tandem lane.

It should be clear that if the TV sub-phase fails on one lane, then a cycle is still lost for the same reasons; and that if a lane fails back to back for both the LV and TV sub-phases, then 2 cycles are lost. Thus, a sub-phase lane failure of either type, regardless of when it happens relative to other failures, wastes exactly one cycle for that lane.

To manage lane failures, we will hold constant the signal sub-phase durations but will reduce the size of the vehicle batches that the pre-signal allows to pass in every cycle. The pre-signal is assumed to be adaptive, normally releasing batches of fixed size, but when a lane has failed the batch size is reduced just enough to ensure that the other lanes receive their full allocation of vehicles. Lane failures can be identified with detector(s) located at the signal stop lines in the tandem lane(s). If during either sub-phase there is any stopped vehicle over the detector(s), the lane fails.

In light of the above, we see that a lane failure does not affect other lanes. Therefore, the long-term discharge rate of any lane can be analyzed independently of the other lanes. In the analysis that follows we will focus on a single lane and evaluate the probability of the lane's failure for the given batch size. Then we shall derive the expected LT and TV flow per cycle for the given batch sizes.

To do this, let $m_.^s$ be the size of the released batch in one of the lanes of the sorting area, where "." denotes L or T. The expected number of vehicles that this lane can discharge in a sub-phase is $m_. = G_./H$. If the coefficient of variation of the saturation headways is γ, then the probability of failure $p_.$ is:

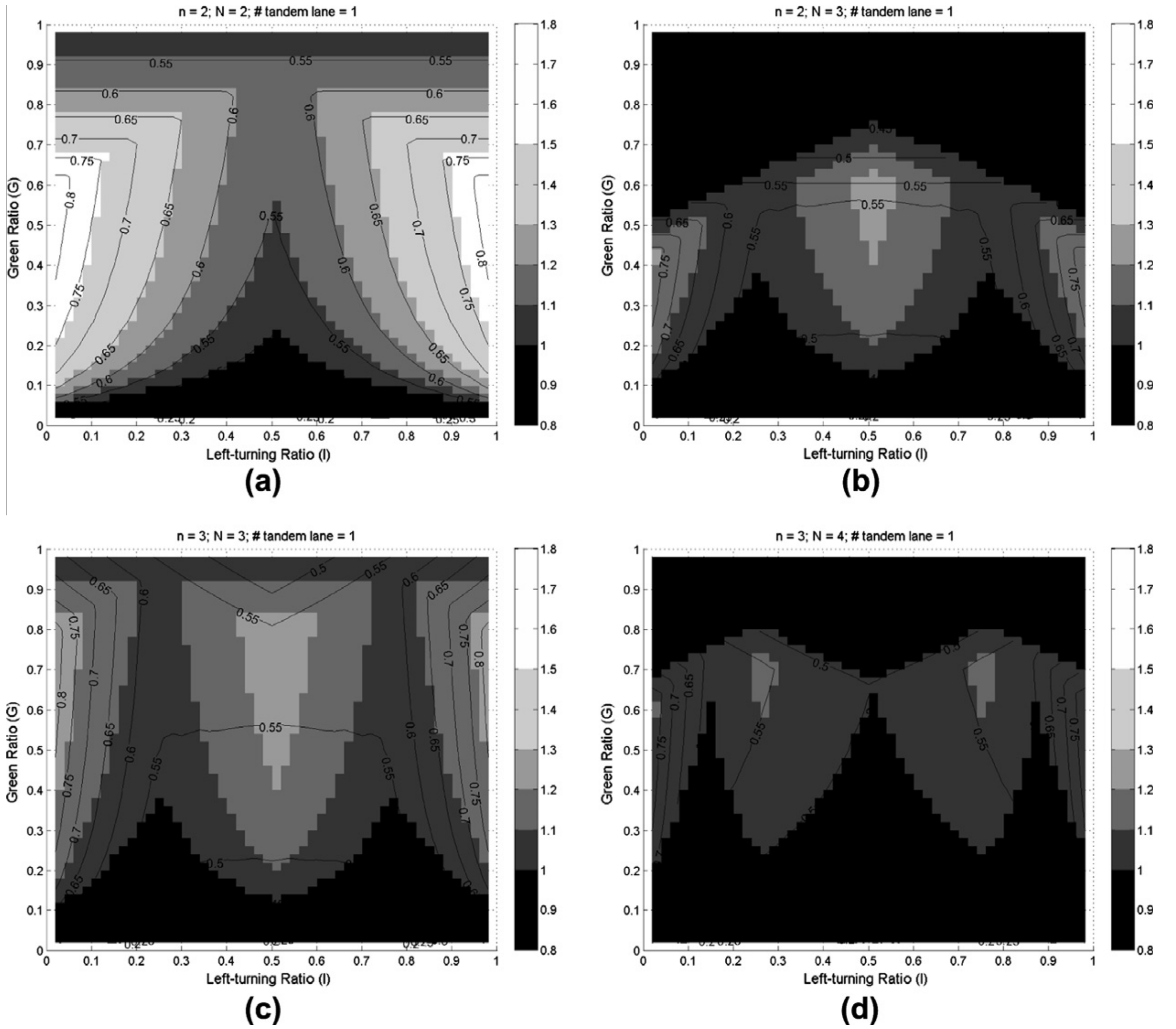

Fig. 4. Stochastic capacities of conventional and tandem design with $C/H = 48$. Contour lines show the ratio of the tandem capacity over the maximal possible flow; shading shows the ratio of the tandem capacity over the conventional capacity: (a) 2 lanes with no turn pocket ($n = N = 2$), 1 tandem lane; (b) 2 lanes with 1-lane turn pocket ($n = 2$, $N = 3$), 1 tandem lane; (c) 3 lanes with no turn pocket ($n = N = 3$), 1 tandem lane; (d) 3 lanes with 1-lane turn pocket ($n = 3$, $N = 4$), 1 tandem lane.

$$p_. = \Pr\left(\sum_{i=1}^{m_.^s} H_i > m_. H\right) = \Phi\left(-(m_. - m_.^s)\Big/\left(\sqrt{m_.^s}\gamma\right)\right)$$

where $\Phi(\cdot)$ is the standard normal c.d.f.[6] This can be written as:

$$p_. = \Phi(-k_.), \quad \text{where } k_. = (m_. - m_.^s)\left(\sqrt{m_.^s}\gamma\right) \quad \text{and} \quad m_. = G_./H \tag{6}$$

The fixed number of vehicles released by the pre-signal in one of its regular cycles (i.e., a cycle without a lane failure) is: $N_L m_L^s + N_T m_T^s$. For any lane, the average number of cycles required to serve its share of these vehicles is $(1 + p_T + p_L)$, because every lane failure implies an unused cycle. Thus, the average flow through the signal is: $\left(N_L m_L^s + N_T m_T^s\right)/[C(1 + p_T + p_L)]$. This can be rewritten in terms of $k_.$ using (6):

$$q^s = \frac{N_L G_L/H\left(1 - k_L\gamma/\sqrt{G_L/H}\right) + N_T G_T/H\left(1 - k_T\gamma/\sqrt{G_T/H}\right)}{C(1 + \Phi(-k_L) + \Phi(-k_T))} \tag{7.1}$$

[6] We assume the saturation headways of different drivers are independently distributed. According to the central limit theorem, $\sum_{i=1}^{m_.^s} H_i$ is approximately normal, since $m_.^s$ is usually large enough.

To express this equation in the special unit system that was used in Section 3.1 without C or H, replace γ (which is the coefficient of variation of the time it takes to discharge one unit of vehicle quantity) by its counterpart γ' in the special unit system. This counterpart is $\gamma' = \gamma/\sqrt{C/H}$ because in the special unit system a unit of vehicle quantity consists of C/H independent vehicles. Thus the equation becomes:

$$q^s = \frac{N_L G_L\left(1 - k_L\gamma'/\sqrt{G_L}\right) + N_T G_T\left(1 - k_T\gamma'/\sqrt{G_T}\right)}{1 + \Phi(-k_L) + \Phi(-k_T)} \quad (7.2)$$

To find the optimum system configuration and signal settings, replace the objective function of TLP with (7.2) and solve the resulting problem. This stochastic version of the design problem will be called STLP. It provides a more realistic assessment of the system than the TLP.[7]

Note the k. variables only appear in the objective function of the STLP. Thus, to solve it, we first maximize (7.1), (7.2) with respect to these variables and then solve for the remaining decision variables. Although the first step is analytically challenging, we find numerically from (7.1) that q^s is a very smooth function of k_L and k_T around the optimum. For example, given $H = 2.5$ s, $G/C = 0.5$, and $\gamma = 0.25$, with C ranging from 30 s to 180 s, and l ranging from 0.1 to 0.9, we find that the value of q^s for $k_L = k_T = 2$ always exceeds 99% of the true optimum.[8] Therefore, we fix these variables at $k_L = k_T = 2$ and then solve the STLP for the remaining variables as in Section 3.1. This is done by evaluating all feasible combinations of N_L and N_T as in Section 3.1.2. The resulting solution should be a conservative approximation of what might be expected in reality. Fig. 4 shows these solutions for the same battery of problems as in Fig. 3, using $C/H = 48$. Note that improvements are still achieved, albeit of a smaller magnitude.

4. Discussion

We close by discussing the benefits of the proposed tandem concept and application issues.

4.1. Potential benefits

Note from Figs. 3 and 4 that low green ratios are problematic under stochastic considerations, even though they were fine in the deterministic cases. This happens because the rate that is produced by the signal in each lane is the product of $(C/H) = 48$ vehicles per lane per cycle and the green ratio; i.e., less than 10 vehicles per lane per cycle if the green ratio is less than 0.2. Such low vehicle numbers magnify the detrimental impact of stochastic fluctuations. However, for more usual green ratios ($0.4 < G < 0.5$), the tandem system with 1 tandem lane increases capacity in all four lane configurations displayed in Fig. 4. The benefit is greater if the intersection is narrow and has no turn pockets. In these cases capacity increases by more than 15% when G is between 0.3 and 0.7, and l is below 0.2. The improvement can exceed 30% for certain combinations of G and l.

Greater benefits are achieved if one allows for more tandem lanes. This is shown in Fig. 5, which evaluates the same intersection configurations as before, but with 2 or 3 tandem lanes. As the figure shows, improvements on the order of 50% can be obtained when $G \approx 0.5$ and the turning ratios are significant.

The reader may complain that the conventional case is unduly penalized in these comparisons because LVs and TVs are not allowed to share lanes in the conventional design—and that a significant penalty would exaggerate the benefits of the tandem configuration. But this penalty turns out to be either non-existent or insignificant. Calculations show that a shared lane actually *reduces* the capacity of the conventional design unless the turning ratio is quite small—and in this case the tandem configuration would not be of much benefit anyway; see Table 1. This table shows critical turning ratios, which if exceeded negate the benefit of a shared lane in the conventional configuration. The reason for the ineffectiveness of shared lanes with a conventional design is that during the LV sub-phase, any TV in a shared lane would block the lane from discharging, and vice versa. The tandem strategy in essence removes the blocking; it even removes the limitation to use only one shared lane. This is why the capacity increases in Fig. 5 are so large.

4.2. Application issues

The tandem concept is not intended for intersections that are always undersaturated, because in this case extra capacity is not needed, and the concept would actually delay some vehicles a little. It should only be applied to reduce or eliminate oversaturation. Ideally, the tandem system would only be activated during the oversaturated times of day, when it can produce a benefit.

To further visualize how effective the concept can be, visit the following url: "http://www.its.berkeley.edu/volvocenter/pre-signal/Tandem_Design.html". This web page displays an animated simulation in CORSIM (FHWA, 2007) of a real 3-lane

[7] The STLP does not include extra constraints in connection with the reduced batch sizes, $m_\cdot^s$, because the deterministic pre-signal green times are always sufficient to accommodate them. To find $m_\cdot^s$, first obtain m. by inserting the optimal G. in the last equality of (6), and then insert m. and the optimum k. in the second equality of (6).

[8] The γ value (coefficient of variation of the saturation headway) of 0.25 is consistent with our field measurement which is 0.24, and the literature: Jin et al. (2009) measured a γ value of 0.3, and Li and Prevedouros (2002) measured a value of 0.22. We find that changing the γ value modestly has no appreciable effect on our predicted outcomes.

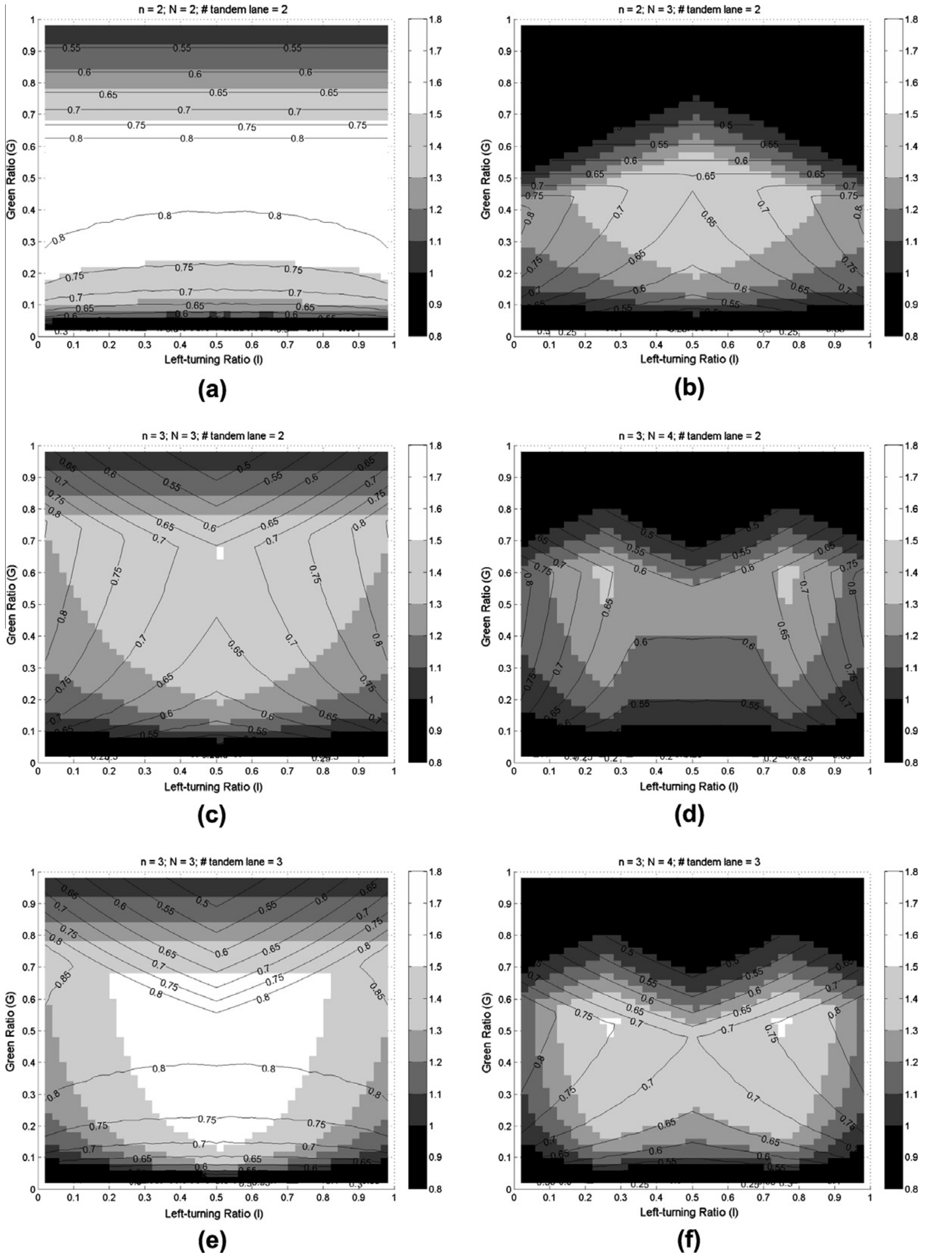

Fig. 5. Stochastic capacities of the conventional and tandem designs with $C/H = 48$. Contour lines show the ratio of the tandem capacity over the maximal possible flow; shading shows the ratio of the tandem capacity over the conventional capacity: (a) 2 lanes with no turn pocket ($n = N = 2$), 2 tandem lanes; (b) 2 lanes with 1-lane turn pocket ($n = 2$, $N = 3$), 2 tandem lanes; (c) 3 lanes with no turn pocket ($n = N = 3$), 2 tandem lanes; (d) 3 lanes with 1-lane turn pocket ($n = 3$, $N = 4$), 2 tandem lanes; (e) 3 lanes with no turn pocket ($n = N = 3$), 3 tandem lanes; (f) 3 lanes with 1-lane turn pocket ($n = 3$, $N = 4$), 3 tandem lanes.

Table 1
Critical left-turning ratios on approaches without turn pockets for a shared lane to be beneficial.

Critical left-turning ratio	$n = N = 2$			$n = N = 3$		
	$C = 60$ s (%)	$C = 90$ s (%)	$C = 120$ s (%)	$C = 60$ s (%)	$C = 90$ s (%)	$C = 120$ s (%)
$G/C = 0.4$	14	8	6	7	4	3
$G/C = 0.5$	11	6	5	5	3	2
$G/C = 0.6$	8	5	4	4	3	2

approach in Chengdu, China, which was studied in Xuan et al. (2010). The approach is currently oversaturated, and this condition cannot be removed by conventional means. The simulation shows side by side how the approach performs under both, the current design and a proposed pre-signal design. While queues grow with the current design, the pre-signal system prevents this growth and serves more vehicles.

Of course, the concept can be used on multiple approaches to an intersection, including on approaches where opposing movements receive green time simultaneously. Though the concept is suitable only for cases when separate phases are given to LVs and TVs on an approach, it can be used when the signal also displays an overlapping phase to simultaneously serve the LVs and TVs on the approach with higher left-turn demand.

The tandem concept requires extra space to hold the transient queues of the vehicles released into the sorting area. Appendix A derives expressions for the lengths of both, the sorting area and the street section upstream of the pre-signal that are required to hold its queues without spill-backs. Notice that this total distance is roughly proportional to cycle length—it turns out to be about 100 m per minute of cycle length. Since this combined distance cannot exceed the approach leg length, short cycles may be needed if city blocks are short. Fortunately, by increasing capacity, pre-signals allow cycles to be shortened. Short cycles can also benefit other users of the intersection, like bicyclists and pedestrians. And even if the distance requirements cannot be met, the concept may still work if suitably modified, though the capacity gains would be lower.

Performance of the tandem concept could be hampered by a lack of driver compliance. For example, an ill-intentioned driver might jump part of the queue in the sorting area by using a pre-signal phase that does not correspond to his desired movement, and then block other drivers while he waits for the right-of-way into the intersection. A video-based enforcement system that records the license plates of offending vehicles and automatically issues citations might minimize the occurrence of this undesirable game-playing.

At this point, the tandem concept is new, and thus we do not know how people will react to it. Field tests are being planned to see how it would work in reality.

Acknowledgements

This work is supported by the National Science Foundation under Grant No. 0856193, and the Volvo Center of Excellence for Future Urban Transport at the University of California, Berkeley.

Appendix A. Spatial considerations

This appendix determines the physical length requirements for the sorting area and the block so that queues do not back up to the pre-signal and to the upstream intersection. We explore a worst case scenario where the tandem design is working at capacity.

First, we derive a formula for the length, D_1, of the sorting area. During each cycle, tandem lanes serve more vehicles than other lanes. Thus the sorting area needs to be just long enough to hold at jam density, K_j, all the vehicles that discharge from a tandem lane each cycle; see time–space diagram in Fig. 6. Clearly then, the minimal length of the sorting area is:

$$D_1 = \left(m_L^s + m_T^s\right)/K_j \quad (8)$$

Next, we express the distances, D_{2L} and D_{2T}, required to hold the queues of the LVs and TVs directly upstream of the pre-signal, assuming that demand equals capacity. These distances should be just long enough to hold at jam density all the vehicles that discharge in each of the pre-signal phases; see Fig. 6. Clearly then:

$$D_{2L} = \frac{C}{H}\frac{g_L^s}{K_j} \quad (9.1)$$

$$D_{2T} = \frac{C}{H}\frac{g_T^s}{K_j} \quad (9.2)$$

The distance that needs to be provided upstream of the pre-signal is therefore:

$$D_2 = \max\{D_{2L}, D_{2T}\} \quad (9.3)$$

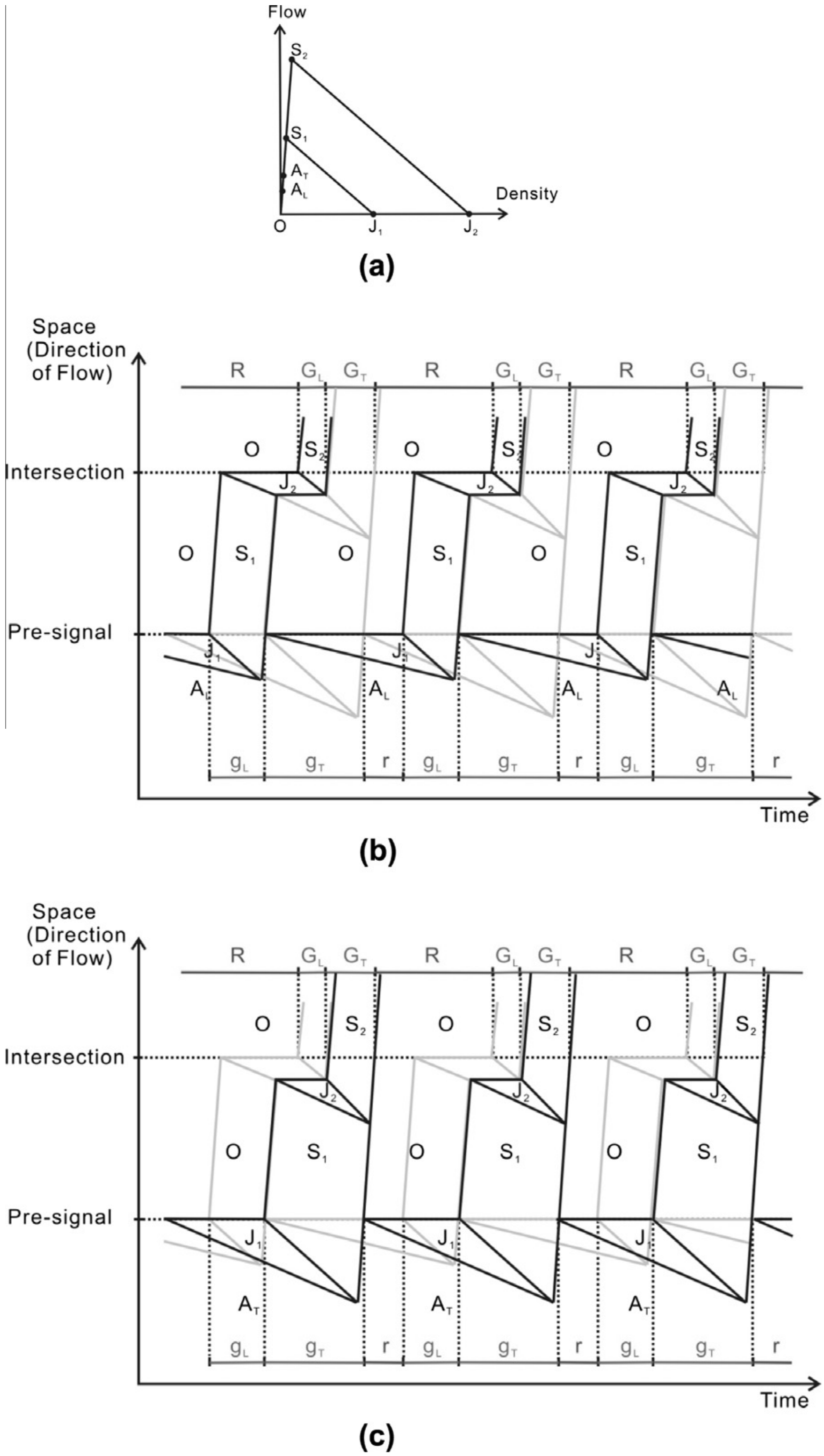

Fig. 6. Spatial evolution during the operation of the tandem design with $n = N_L = N_T = N = 2$ and saturated demand: (a) fundamental diagram; (b) time–space diagram of the LVs (with that for TVs in gray color); (c) time–space diagram of the TVs (with that for LVs in gray color).

References

Allsop, R.E., 1972. Estimating the traffic capacity of a signalized road junction. Transportation Research 6 (3), 245–255.

Al-Salman, H.S.T., Salter, R.J., 1974. The control of right turning vehicles at signal controlled intersections. Traffic Engineering & Control 15 (15), 683–686.

ATTAP, University of Maryland, 2006. Unconventional Arterial Intersection Design. <http://attap.umd.edu/uaid.php>.

Daganzo, C.F., 2007. Urban gridlock: macroscopic modeling and mitigation approaches. Transportation Research Part B 41 (1), 49–62.

Federal Highway Administration, 2007. Traffic Analysis Toolbox Volume IV: Guidelines for Applying CORSIM Microsimulation Modeling Software, No. FHWA-HOP-07-079. <http://ops.fhwa.dot.gov/trafficanalysistools/tat_vol4/vol4_guidelines.pdf>.

Gallivan, S., Heydecker, B.G., 1988. Optimising the control performance of traffic signals at a single junction. Transportation Research Part B 22 (5), 357–370.

Greenough, J., Kelman, L., 1999. ITS technology meeting municipal needs – the Toronto experience. In: 6th World Congress Conference on ITS. Toronto, Canada, 8–12 November.

Goldblatt, R., Mier, F., Friedman, J., 1994. Continuous flow intersection. Institute of Transportation Engineers Journal 64 (7), 35–42.

Improta, G., Cantarella, G.E., 1984. Control system design for an individual signalized junction. Transportation Research Part B 18 (2), 147–167.

Jin, X., Zhang, Y., Wang, F., Li, L., Yao, D., Su, Y., Wei, Z., 2009. Departure headways at signalized intersections: a log-normal distribution model approach. Transportation Research Part C 17 (3), 318–327.

Li, H., Prevedouros, P.D., 2002. Detailed observations of saturation headways and start-up lost times. Transportation Research Record 1802, 44–53.

Lo, H.K., Chow, A.H.F., 2004. Control strategies for oversaturated traffic. Journal of Transportation Engineering 130 (4), 466–478.

Newell, G.F., 1989. Theory of Highway Traffic Signals. Research Report, Institute of Transportation Studies. University of California, Berkeley, UCB-ITS-RR-89-7

Oakes, J., ThellMann, A.M., Kelly, I.T., 1994. Innovative bus priority measures. In: Proceedings of Seminar J, Traffic Management and Road Safety, 22nd PTRC European Transport Summer Annual Meeting, vol. 381, pp. 301–312.

Peck, C., Blanco, J., Lopez, J., 1999. Learning from the User – Next Steps for ITACA's Adaptive Control. Traffic Technology International, Annual Review, pp. 155–158.

Rodegerdts, L.A., Nevers, B., Robinson, B., 2004. Signalized Intersections: Informational Guide. No. FHWA-HRT-04-091. <http://www.tfhrc.gov/safety/pubs/04091/04091.pdf>.

Silcock, J.P., 1997. Designing signal-controlled junctions for group-based operation. Transportation Research Part A 31 (2), 157–173.

Sussman, J., et al., 2000. What have we Learned about Intelligent Transportation Systems. No. FHWA-OP-01-006. <http://www.itsdocs.fhwa.dot.gov/JPODOCS/REPTS_TE/13316.pdf>.

Webster, F.V., 1958. Traffic Signal Settings. Road Research Lab Technical Paper No. 39.

Wong, C.K., Wong, S.C., 2003. Lane-based optimization of signal timings for isolated junctions. Transportation Research Part B 37 (1), 63–84.

Wu, J., Hounsell, N., 1998. Bus priority using pre-signals. Transportation Research Part A 32 (8), 563–583.

Xuan, Y., Gayah, V., Daganzo, C.F., Cassidy, M.J., 2009. Multimodal Traffic at Isolated Signalized Intersections: New Management Strategies to Increase Capacity. Working Paper, Volvo Center of Excellence for Future Urban Transport, University of California, Berkeley, UCB-ITS-VWP-2009–10.

Xuan, Y., Li, Y., Daganzo, C.F., Cassidy, M.J., 2010. Improving Flows of Buses and Cars through an Intersection by Enforcing Bus-only Lane and Increasing Car Flow Capacity with Pre-signal. Volvo Technical Report (in Chinese). <http://www.its.berkeley.edu/volvocenter/pre-signal/Chengdu_TechnicalReport_chn.doc>.