TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
2022, Vol. 2676(4) 649–668

DOI: 10.1177/03611981211064281
journals.sagepub.com/home/trr
SAGE

Incorporating Delay Minimization in Design of the Optimized Arterial Signal Progression

Yen-Hsiang Chen[1], Yao Cheng[1], and Gang-Len Chang[1]

Abstract

Despite the abundance of studies on signal progression for arterial roads, most existing models for bandwidth maximization cannot concurrently ensure that the resulting delays will be at a desirable level, especially for urban arterials accommodating high turning volume at some major intersections or constrained by limited turning bay length. Extending from those models that aim to address delay minimization in the progression design, this study provides two enhanced progression maximization models for arterials with high turning volumes. The first model aims to select the signal plan that can produce the lowest total signal delays for all movements from the set of non-inferior offsets produced by MAXBAND. Failing to address the impact of potential turning bay spillback at some critical intersections under such a design may significantly degrade the quality of through progression and increase the overall delay. For this reason, the second model proposed in this study offers the flexibility to trade the progression bandwidths within a pre-specified level for the target delay reduction, especially for turning traffic. The evaluation results from both numerical analyses and simulation experiments have shown that both proposed models can produce the desirable level of performance when compared with the two benchmark models, MAXBAND and TRANSYT 16. The second model yielded the lowest average network delay of 117.2 seconds per vehicle (s/veh), compared with 121.7 s/veh with TRANSYT. Moreover, even its average delay of 141.8 s/veh for through vehicles is comparable with that of 141.2 s/veh by MAXBAND, which is designed mainly to benefit through-traffic flows.

Keywords

operations, traffic control devices, signalized, traffic signals, traffic signal systems, signal phase, signalized intersection, traffic signal

To improve the efficiency of urban arterial roads accommodating heavy traffic volumes, numerous methods for signal optimization have been developed by the transportation community over the past several decades. Depending on the control objectives, most existing models belong to one of two categories: maximizing signal progression or minimizing delay. Although those models in the former category with the produced progression bands and with time-space diagrams are preferred in practice, there are some critical areas for potential enhancements. First, an arterial designed with the objective of progression band maximization, because of the inherited mathematical properties, can actually have multiple sets of offsets to yield the same total bandwidth but different total vehicle delays. Secondly, since most progression-based models mainly focus on efficiently progressing through-traffic flows, the resulting delays from the perspective of the entire arterial may exceed desirable levels, especially for arterials experiencing significant turning flows or having insufficient turning bay lengths at some major intersections. The excessive delay caused to the large volume of turning flows under the progression-for-through-traffic design may degrade the performance of the arterial control, and may even cause turning bay spillbacks and partially block the through traffic in the progression bands.

Figure 1 shows an example of MAXBAND's solution (*1*) on a time-space diagram with three illustrative cases

[1]Department of Civil and Environmental Engineering, University of Maryland, College Park, MD

Corresponding Author:
Yao Cheng, chengyao09@gmail.com

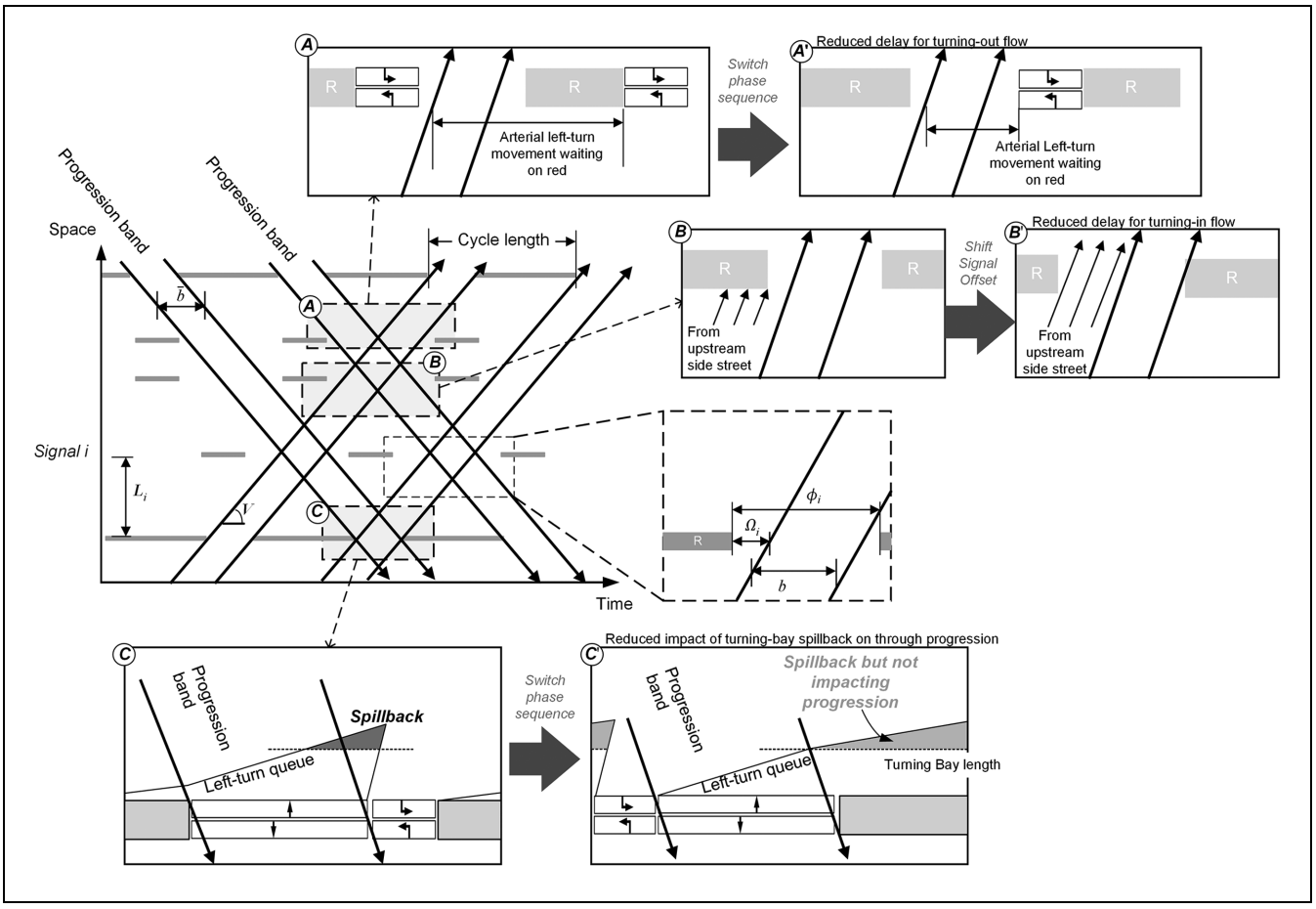

Figure 1. Key notations in MAXBAND and its multiple optimal solutions with the same bandwidths but various impacts on the total delay.

where three alternative signal plans (see areas *A*', *B*', and *C*') can produce the same maximum bandwidths but different total delays over the arterial. When comparing scenarios *A* and *A*' in Figure 1, one may notice that the signal delay of left-turn vehicles from the arterial can be significantly reduced by adjusting the phase sequences and offsets without affecting the progression bandwidth. Such a delay reduction can also be provided to turning-in flows from the side streets, as evidenced in the results of comparison between scenarios *B* and *B*'.

It should be noted that, among those non-inferior optimal signal plans produced from the bandwidth maximization models, some can yield much smaller total vehicle delays, because of the lower likelihood of left-turn bay overflows impeding the through progression traffic. For example, as shown in scenarios C and C', proper adjustments of the phase sequences from the latter can yield a reduction in both the delay and the impacts from potential turning bay overflows with the same progression bandwidth.

Based on the potential enhancements highlighted in Figure 1, this study presents two enhanced arterial signal progression models. The proposed models, extending the notion of MAXBAND, are mainly for arterials accommodating considerable turning volumes or bay length constraints, offering the flexibility to trade the progression bandwidth from its optimal state with a significant overall delay reduction. Their embedded formulations are designed to address the following key issues:

- How to estimate the queuing delay of both through and turning flows under a signal progression plan;
- How to identify the likelihood of turning bay spillbacks occurring and estimate the resulting impact on the through progression flows;
- How to identify the signal plan that can yield the lowest total delay and least impact from left-turn bay spillback out of the multiple non-inferior solutions generated from the MAXBAND-based arterial progression models; and
- How to specify an acceptable bandwidth tradeoff that can concurrently benefit turning movements and ensure sufficient progression for through movements.

Literature Review

A review of related literature indicates that maximizing two-way progression was first investigated systematically by Morgan and Little (*2, 3*) to offer through vehicles from both inbound and outbound directions opportunities to travel along the entire arterial without stopping. MAXBAND, later proposed by Little et al. (*1*), formulated a progression band maximization problem with a mixed-integer linear programming (MILP) model that can be effectively solved with commercial packages. Along the same line, MULTIBAND (*4*) and PASSER (*5*) allow the arterial to design with variable bandwidths based on the volumes on each link. Recently, with the potentially available information on path-flow or origin–destination (OD), a set of progression bands for selected paths or particular OD flows can be implemented with the multi-path model (*6*), OD band (*7*) or other models following the same notion, such as Yan et al. (*8*). Such concepts have also been extended to the network level and unconventional intersections (*9–12*).

One promising family of models employing delay minimization as the objective function is TRANSYT (*13*), which relies on an embedded platoon dispersion equation to reflect the interrelationship between a link's upstream and downstream flows. With a hill-climb optimization method to search for a lower delay at each iteration, it can generate offline the signal plans of minimal delay for large-size networks. Shockwave theory is also another popular methodology for delay computation, especially for oversaturated intersections (*14*). Grounded in the cell transmission methodology, Lo et al. (*15*) provided a delay-minimizing model for an urban arterial by computing the differences between the number of vehicles traveling with free-flow speed and those held in cells. Stevenovic et al. (*16*) applied a framework with an embedded microscopic model to calculate the delay and optimized the network signal plan with genetic algorithms. Several other prominent models by different researchers—for example, Kashani and Saridis (*17*), Yun and Park (*18*), and Liu and Chang (*19*)—have also been developed along the same line.

Recognizing the desire to have an arterial signal control plan with strengths from the models in both categories, some studies have attempted to optimize those two objectives either sequentially or concurrently. For example, Wallace and Courage (*20*) proposed an alternative objective function to maximize "PROS/PI," the ratio of progression opportunity over the disutility performance indices in TRANSYT to provide an arterial progression plan while minimizing the delays and number of stops. Acknowledging that the final optimal solution from TRANSYT varies with the specified initial solutions, Cohen (*21*) applied the results from MAXBAND as the initial solution such that it is more likely to produce the progression bands. Cohen and Liu (*22*) further proposed the bandwidth-constrained (BWC) optimization for TRANSYT to ensure that the improvement in delay will not reduce the bandwidths over the optimization process.

For the same purpose but with a different methodology, Chang et al. (*23*) and Chang and Messer (*24*), recognizing the existence of multiple optimal offsets and slackness in multiple non-critical intersections from MAXBAND's solution, proposed to adjust relative offsets at each set of neighboring intersections based on empirical relations between offsets and delay. The study by Lan et al. (*25*) applied a piecewise linear function to formulate the approximated nonlinear delay in their proposed MILP model for signal progression. Focusing on reducing the delay for left-turning vehicles in an arterial's progression design, Chen et al. (*26*) formulated a two-stage model to offer left-turn flows with the highest "left-turn efficiency index" to approximate the minimized delay. Addressing similar concerns, but mainly accounting for the impact of turning-in vehicles from side streets on the progression efficiency of through traffic, Chen et al. (*27*) presented a methodology to provide local progression bandwidth for such flows.

Noticeably, the delays addressed in the aforementioned hybrid models are mainly for queuing vehicles during the red phases. However, the potential turning bay overflows and their impacts on through progression have not yet been addressed.

Model Formulations

In light of the aforementioned enhancement needs, this study, grounded in the core logic of MAXBAND (*1*), proposes two arterial progression models that intend to minimize the intersection delays with the optimized offsets and phase sequences. Model 1 aims to select the signal plan that produces the lowest delays among the set of multiple optimal solutions by MAXBAND. Recognizing that failing to address the issue of turning bay spillback at some critical intersections may significantly degrade the quality of progression and substantially reduce the arterial's "effective" through bandwidth, this study has further proposed Model 2 to concurrently minimize the impact of left-turn bay spillbacks, including its resulting delays, by allowing a marginal reduction of the progression bandwidth from its theoretical optimal level.

The Objective Function of Model 1

Aiming to select the signal progression plan producing the lowest total delay among the multiple optimal solutions by MAXBAND, the objective of Model 1, as expressed in Equation 1, includes delays from through

and left-turn queue evolutions during the red phases and excessive delay caused by left-turn bay overflows.

$$\text{Min}\, D^R + D^S \tag{1}$$

where D^R is the delay from vehicle queues during the red phase and D^S is the excessive delay of through vehicles caused by turning bay spillbacks. Note that the optimal signal plan for Equation 1 should be subject to the bandwidth constraints, shown in Equations 2 and 3, to ensure the highest possible progression efficiency for through vehicles.

$$b = b^* \tag{2}$$

$$\bar{b} = \bar{b}^* \tag{3}$$

where b^* and $\bar{b}^*$ denote the maximum inbound and outbound bandwidths obtained from MAXBAND.

With the objective function and Equations 2 and 3, this model was designed with the capability of finding a solution denoted by area-A ', area-B ', and area-C ' (see Figure 1) to reduce the delay during the red phase (D^R), and the delay caused by turning bay overflow (D^S).

The Objective Function of Model 2

The objective function of Model 2, focusing on ensuring the through progression band not interrupted by turning bay spillbacks, can be expressed with Equation 4.

$$\text{Min}\ W^S \cdot N^S + D^R + D^S \tag{4}$$

where N^S denotes the total number of intersection approaches with turning bay spillback and W^S is the weighting factor. Note that because of the serious impacts of turning bay spillbacks on the progression effectiveness, Model 2's first priority is to minimize the number of turning bays that incur overflows. As such, its objective function in Equation 4 has assigned a weighting factor, W^S, to each intersection with a numerical value significantly greater than the two other terms to reflect the proposed control system's priority.

Note that a signal plan, as stated previously, yielding the maximal progression bandwidths for through vehicles may encounter overflows from some critical left-turn bays and thus incur excessive delays from potential impedance or partial blockage of the through lanes by the spilling queues. Conceivably, reducing the bandwidth from its maximal level offers room for further reduction of the queuing delay, and the probability of traffic overflows by refining the offsets and the phase sequence (see the signal plan highlighted in area-C' of Figure 1). This study has thus developed Model 2 for trading the marginal bandwidth reduction with a substantial delay decrease, and also the likelihood of turning bay overflows at major intersections. The allowable bandwidth reduction, as shown in Equations 5 and 6, can be captured by introducing a parameter, denoted as p, to indicate the allowable level for trading with the delay reduction.

$$b^* \geq b \geq (1-p)b^* \tag{5}$$

$$\bar{b}^* \geq \bar{b} \geq (1-p)\bar{b}^* \tag{6}$$

For example, by setting $p = 5\%$, the refined signal plan for delay minimization must be constrained to preserve 95% of the maximum bandwidth. As shown in Equation 7, the critical parameter p should be calculated based on the maximum bandwidths obtained from MAXBAND and the time duration needed to dissipate the longest initial queue along the arterial.

$$p = \left(1 - \frac{t_d}{b_m}\right) \cdot 100\% \tag{7}$$

where t_d denotes the dissipation time of the longest initial queue among all intersections along the arterial; and b_m denotes the maximum bandwidth obtained from MAXBAND.

Under such an objective function and related constraints, queue impedance at critical intersections can be minimized to substantially reduce the arterial's overall delay, despite such changes possibly producing a marginally reduced bandwidth for the through flows.

Model Framework

Figure 2 shows the framework of the proposed models with the above objectives, along with the embedded formulations designed to perform the following tasks: (i) computing the two-way progression bandwidths and specifying their upper and lower bounds based on the results from MAXBAND; (ii) estimating the queues and the resulting delay for the arterial's through and left-turn vehicles with their arrival patterns and the queue evolving dynamics; and (iii) identifying the set of left-turn bays that may experience overflows and estimating the resulting delays. The notations for key model variables and parameters are listed in Table 1.

Design of the Two-Way Progression

To formulate the two-way progression bands over the arterial, one can follow the core logic of MAXBAND (*1*) and apply the following *interference constraints* to ensure that the band, as shown in Figure 1, would only be within the green interval:

$$\Omega_i + b \leq \phi_i \forall i \in \mathfrak{l} \tag{8}$$

$$\Omega_i + \bar{b} \leq \phi_i \forall i \in \mathcal{O} \tag{9}$$

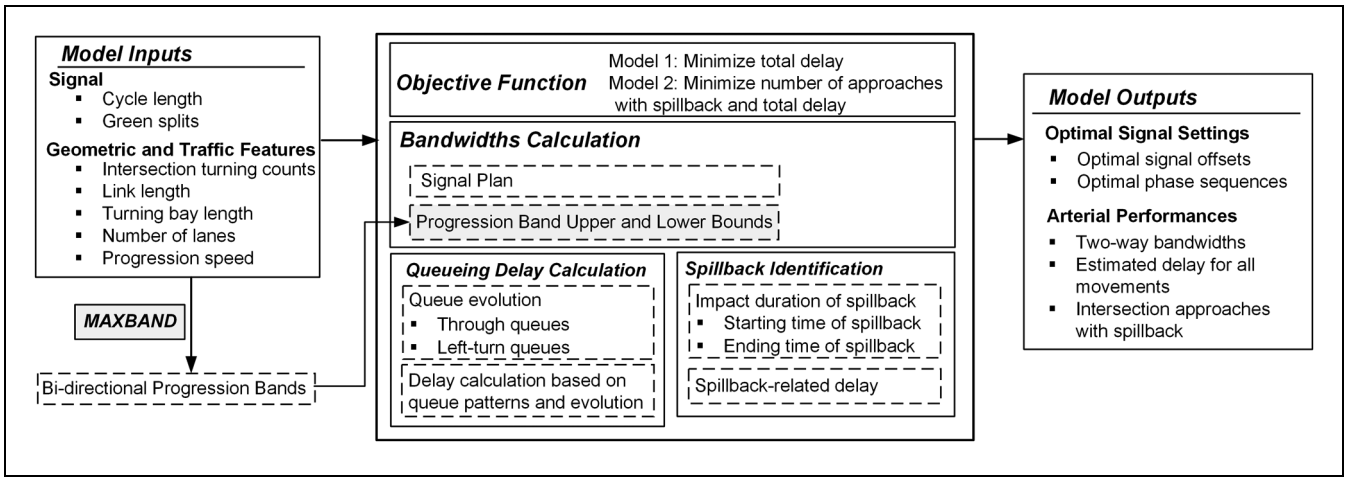

Figure 2. Framework of the proposed model.

where ϕ_i is the green duration for the through movement at signal i; b ($\bar{b}$) is the through bandwidth for inbound(outbound) direction; Ω_i is the starting time lag between the green interval for the through movement and the progression band at the downstream intersection of link i; and $\mathfrak{l}$ ($\mathcal{O}$) is the set of inbound (outbound) links.

Loop integer constraints (*1*) in MAXBAND are then introduced to ensure that the band will exist between any pair of two adjacent intersections, as shown below:

$$\theta_i + \Omega_i = \theta_{i-1} + \Omega_{i-1} + \frac{L_i}{V} + n_{i-1} \cdot c \qquad (10)$$

where θ_i is the start time of a green interval at signal i; L_i denotes the link length; V is the cruising speed; c is cycle length; and n_i is an integer for ensuring that both sides of the equation refer to the same signal cycle. Note that Equation 10 applies to both inbound and outbound directions. The formulations hereafter will be introduced for the inbound direction, and those for the outbound direction can be developed with the same logic.

Considering the impact of various left-turn phase patterns on the arterial's bandwidths and delays, Equations 11 and 12 are developed to facilitate the design of variable phase sequences:

$$\theta_i = \Theta_j + \chi_{O(i)} \cdot (\phi^L_{O(i)} + I) \qquad (11)$$

$$\theta^L_i = \Theta_j + (1 - \chi_i) \cdot (\phi_{O(i)} + I) \qquad (12)$$

where χ_i is a binary variable indicating the leading inbound left-turn signals at downstream of link i; Θ_j is the intersection offset; θ^L_i is the starting time of the green interval for the left-turn movement; ϕ_i (ϕ^L_i) is the duration of the green interval for the through (left-turn) movement; $O(i)$ is the outbound signal for the opposite direction at intersection i; and I is the inter-green duration.

Queue Evolution for Through Vehicles

Time-dependent queue evolution is essential to evaluate the queuing delay and compute the starting and ending times of turning bay spillbacks, if they occur. One of the most efficient ways to reflect dynamic queue evolution at intersections is to adopt the shockwave theory (*28–34*). Figure 3a illustrates a typical queue evolution process for through vehicles, consisting of three queue formation shockwaves. Wave 1 and Wave 3, respectively, denote the queue formation shockwaves formed by arriving through vehicles, discharged before and after the upstream intersection's red phase, but stopped by the red phase at the subject intersection. Wave 2 refers to those queues contributed by vehicles from the upstream intersection's side street (i.e., during the red phase for through vehicles).

Note that the queuing delays, as shown in Figure 3*a*, can be represented with the gray area, between *queue formation* and *discharging shockwaves*, in the time-space diagram. This section will first present the formulations to capture the arrival pattern of through vehicles over the arterial link between two coordinated neighboring signals, and then characterize their temporal and spatial queues.

By denoting w^1_i, w^2_i, and w^3_i, respectively, as the backward propagating speeds of queue formation shockwaves for Waves 1, 2, and 3 (shown in bold lines in Figure 3*a*), one can represent their time-varying queues as follows (see Figure 3*b*).

$$Q_i(t) = \begin{cases} w^1_i t & 0 \le t \le y^1_i \\ Q^1_i + w^2_i \cdot (t - y^1_i) & y^1_i \le t \le y^2_i \\ Q^2_i + w^3_i \cdot (t - y^2_i) & y^2_i \le t \le \sigma_i \end{cases} \qquad (13)$$

where $Q_i(t)$ is the time-varying location of the through queues' tail (in feet) at intersection i; y^1_i is the

Table 1. Key Notations in the Proposed Formulation

Indices, sets, and operations	
i	Link index
j	Intersection index
$O(i)$	The link at the opposite direction facing the same intersection with link i
$\mathcal{I}$	The set of links with traffic flow heading inbound
$\mathcal{O}$	The set of links with traffic flow heading outbound
Parameters	
b^* ($\bar{b}^*$)	Maximum attainable bandwidth of inbound(outbound) traffic (s/cycle)
W^S	The weighting factor for setting the priority of spillback minimization (–)
p	Allowable tradeoff percentage of progression bandwidths (%)
V	Progression speed (ft/s)
c	Common cycle length (s)
ϕ_i ($\phi^L{}_i$)	Green split of through (left-turn) movement departing from link i (s)
r_i ($r^L{}_i$)	Red phase of through (left-turn) movement departing from link i (s)
I	Inter-green duration (s)
L_i	Length of link i (ft)
L_i^B	Left-turn bay length of link i (ft)
w_i^ι (w_i^κ)	Backward propagation speed of queue formation shockwave of through (left-turn) lanes of link i when Wave ι (κ) prevails (ft/sec), $\iota \in \{1,2,3\}$; $\kappa \in \{4,5,6\}$
z (z^L)	Backward propagation speed of discharging shockwave (ft/sec) of through (left-turn) lanes (ft/s)
N_i (N_i^L)	Number of through (left-turn) lanes on link i (–)
f_{i-1}^T	Arrival flowrate to link i from upstream through movement that will be impacted if left-turn spillback occurs (veh/s)
k^j	The jam density (veh/ft)
Variables	
b ($\bar{b}$)	Progression bandwidth of inbound(outbound) direction (s)
N^S	Integer variable to count the number of intersection approaches with turning bay spillback (−)
S_i	Binary indicator equaling one if through movement of link i is affected by left-turn spillback and zero otherwise (−)
θ_i (θ_i^L)	Starting time of green interval of through (left-turn) movement of link i (s)
Θ_j	Offset of intersection j (s)
χ_i	Binary variable showing one if with leading left-turn phase and zero otherwise (−)
Ω_i	The time differences between the starting time of through green interval and the starting time of progression bandwidth passing link i (s)
d_i ($d^L{}_i$)	Delay of through (left-turn) movements queue on the red phase of link i (veh-s/cycle)
$d^S{}_i$	Delay of through movements on link i caused by left-turn bay spillback (veh-s/cycle)
d^X ($d^{L,X}$)	Delay of through (left-turn) movements with arrival pattern X (veh-s/cycle), X ∈ {A, B, C}
d^D	Delay of left-turn movements with arrival pattern D (veh-s/cycle)
δ_i^X ($\delta_i^{L,X}$)	Binary indicator which equals one if arrival pattern X (X ∈ {A, B, C}) holds for through (left-turn) movement and zero otherwise (−)
δ_i^D	Binary indicator which equals one if arrival pattern D holds for left-turn movement and zero otherwise (−)
y_i^1 ($y_i^{L,1}$)	Arrival time of the first vehicle in a cycle from upstream side street joining queue of through (left-turn) movement of link i if arrival pattern A, B, or C applies (s)
y_i^2 ($y_i^{L,2}$)	Arrival time of the last vehicle in a cycle from upstream side street joining queue of through (left-turn) movement of link i if arrival pattern A, B, or C applies (s)
σ_i (σ_i^L)	Timepoint of *physical queue* being fully discharged for through (left-turn) movement of link i (s)
Q_i^1 ($Q_i^{L,1}$)	Queue length when the first vehicle in a cycle from upstream side street joining queue of through (left-turn) movement of link i if arrival pattern A, B, or C applies (ft)
Q_i^2 ($Q_i^{L,2}$)	Queue length when the last vehicle in a cycle from upstream side street joining queue of through (left-turn) movement of link i if arrival pattern A, B, or C applies (ft)
Q_i^M ($Q_i^{L,M}$)	Maximum physical queue length of through (left-turn) movement at link i (ft)
$\hat{y}_i^{L,1}$ ($\hat{y}_i^{L,2}$)	Arrival time of the first (last) vehicle in a cycle from upstream through movement joining the queue of left-turn movement of link i if arrival pattern D applies (s)
$\hat{Q}_i^{L,1}$ ($\hat{Q}_i^{L,2}$)	Queue length (ft) when the first (last) vehicle in a cycle from upstream through movement stops joining queue of through (left-turn) movement of link i if arrival pattern D applies (ft)
π_i^4 (π_i^6)	Duration of impact on through movements arriving from upstream from left-turn spillback when Wave 4 (6) prevails for left-turn traffic (s)
π_i^D	Duration of impact on through movements arriving from upstream from left-turn spillback when left-turn arrival pattern is D (s)
n_i	Integer variable to match to the same cycle for link i and its downstream (−)
δ_i^S	Binary variable equaling one if turning bay spillback impacting through movements on link i and zero otherwise

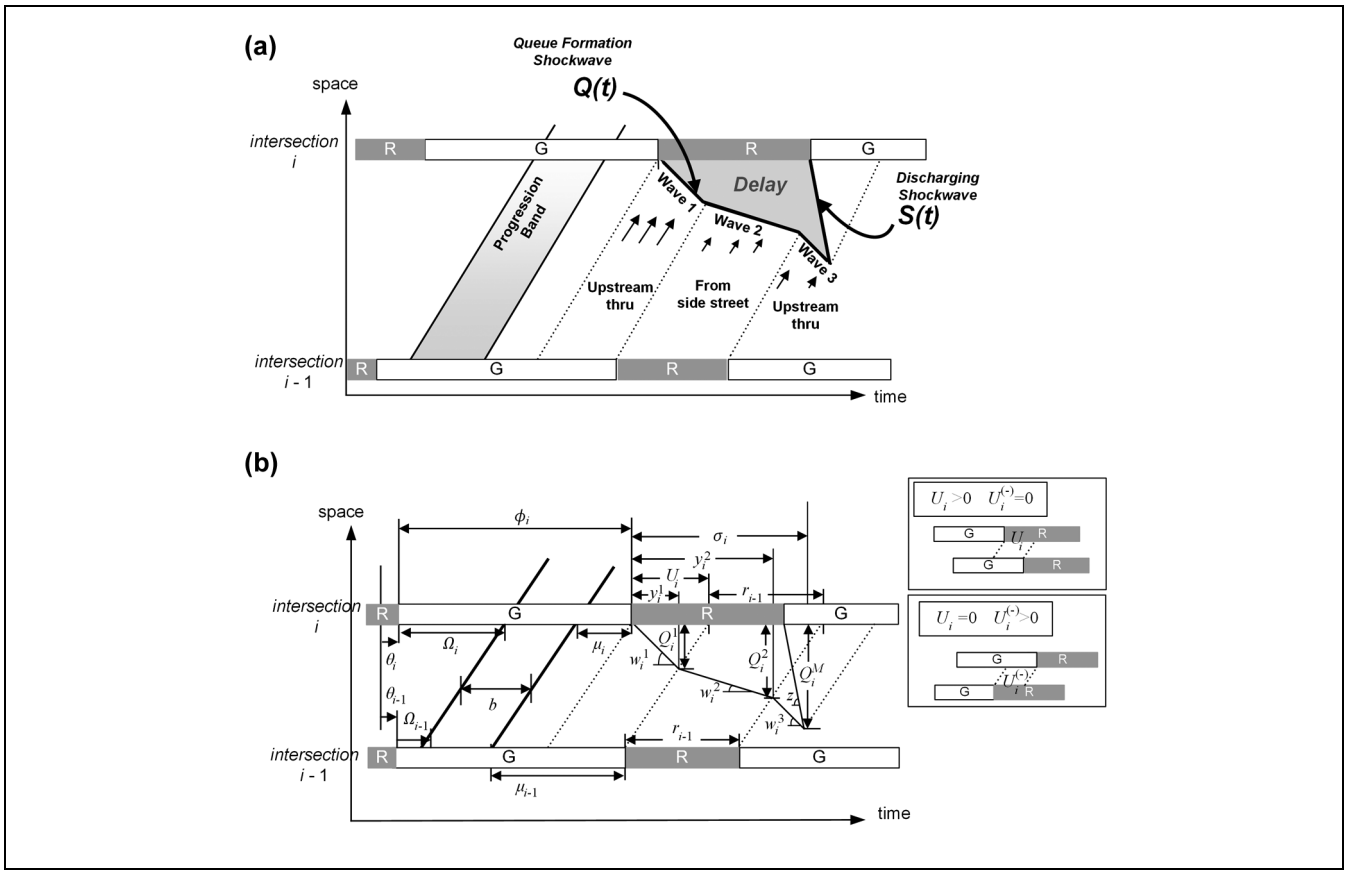

Figure 3. Evolution of the traffic queues and shockwaves between two intersections: (*a*) the tail of through queues with Waves 1 to 3 and (*b*) illustration of key variables.

time point when the last vehicle in a cycle from the upstream through phase encounters the tail of through queues at the location Q_i^1; y_i^2 is the time point when the last vehicle in a cycle from the upstream side streets joins the through queues at the location Q_i^2; and σ_i is the queue dissipation time. Note that all those time-dependent variables associated with the evolution of queues discussed hereafter are referenced from the onset of the red phase.

Depending on an arterial's signal coordination design, each link may not experience all three types of queue formation shockwave. Focusing on the evolution of the time-dependent queues (*35*), Figure 4 shows the following three primary arrival patterns contributing to the physical queue formation:

Pattern A- The through queue consists of solely the upstream through vehicles;
Pattern B- Discharging shockwave catches up the accumulating queue with Wave 2; and
Pattern C- Discharging shockwave catches up the accumulating queue with Wave 3.

By taking into account various arrival patterns, one can then specify Equation 14, a refined version of Equation 13, to approximate various queue lengths under those three shockwaves:

$$Q_i(t) = \begin{cases} w_i^1 t & 0 \le t \le \min(y_i^1, \sigma_i) \\ \min[Q_i^1 + w_i^2 \cdot (t - y_i^1), Q_i^M] & \min(y_i^1, \sigma_i) \le t \le \min(y_i^2, \sigma_i) \\ \min[Q_i^2 + w_i^3 \cdot (t - y_i^2), Q_i^M] & \min(y_i^2, \sigma_i) \le t \le \sigma_i \end{cases} \quad (14)$$

where Q^M is the maximum queue length constrained by the link length, as shown in Equation 15, which can be computed with Equation 16.

$$Q_i^M \le L_i \quad (15)$$

$$Q_i^M = (\sigma_i - r_i)z \quad (16)$$

where L_i is the link length, r_i is the red phase duration, and z is the backward speed of the *discharging shockwave*.

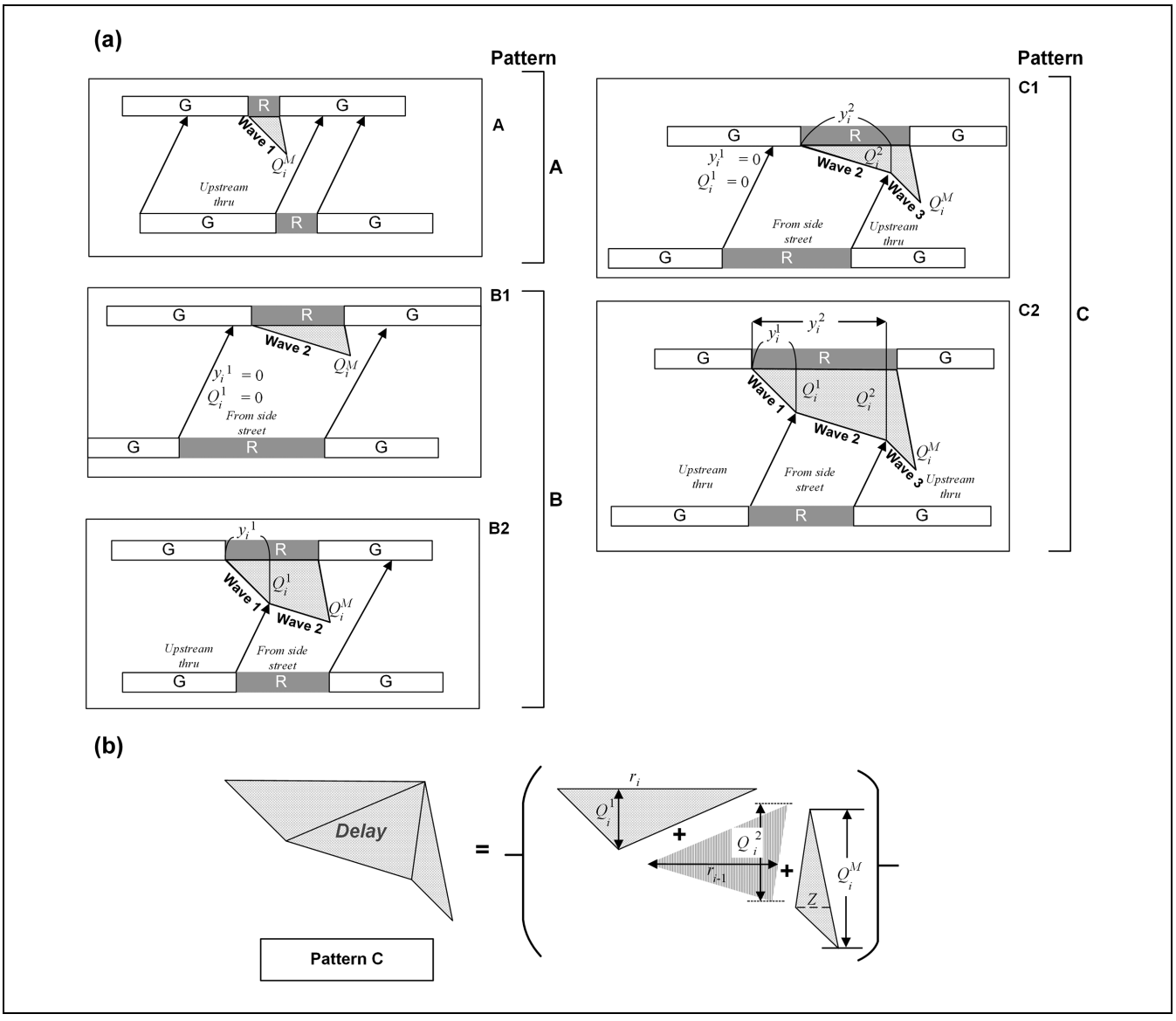

Figure 4. Computing the delay based on the arrival patterns and queue evolution: (*a*) possible arrival patterns and queue evolutions based on the signal offsets between neighboring intersections (*35*) and (*b*) through queuing delay with arrival Pattern C.

Delay Estimation for Through Vehicles

With the above queue evolution process, one can then estimate the queuing delay of the through vehicles based on the queue patterns actually taking place on each arterial link. Since Pattern C evolution comprises Waves 1 to 3, the presentation hereafter will focus on detailing its resulting delay; the same computation can be done for other patterns.

As shown in Figure 4*b*, the through queue delay under Pattern C can be calculated by summing up the areas of three shaded triangles reflecting the queue evolution process, as follows:

$$d_i^C = (k^j \cdot N_i/2)(r_i Q_i^1 + r_{i-1} Q_i^2 + Z Q_i^M) \quad (17)$$

$$Z = \frac{Q_i^2 + (r_i - y_i^2)z}{z} \quad (18)$$

where d_i^C (s-veh) is the total through delay on link i when queue pattern C prevails, Z is the duration shown in Figure 4*b*, representing the time point when the *queue discharging shockwave* reaches the location of Q_i^2, measured from y_i^2, and k^j is jam density. These three terms in Equation 17, $r_i Q_i^1$, $r_{i-1} Q_i^2$, and $Z Q_i^M$, denote, respectively, the three shaded triangle areas shown in Figure 4*b*.

The delay formula for patterns A and B, shown in Table 2, can be developed with the same notion.

Table 2. Delay Computation on Red Phase of Arrival Patterns

Arrival pattern	Delay formulations
Through movement	
A	$d_i^A = (k^j \cdot N_i/2)\, zQ_i^M$
B	$d_i^B = (k^j \cdot N_i/2)(r_i Q_i^1 + XQ^M),\ X = \frac{Q_i^1 + (r_i - y_i^1)z}{z}$
C	$d_i^C = (k^j \cdot N_i/2)(r_i Q_i^1 + r_{i-1} Q_i^2 + ZQ_i^M),\ Z = \frac{Q_i^2 + (r_i - y_i^2)z}{z}$
Left-turn movement	
$\mathbf{A_L}$	$d_i^{L,A} = (k^j \cdot N_i^L/2)\, z \cdot Q^{L,M}$
$\mathbf{B_L}$	$d_i^{L,B} = (k^j \cdot N_i^L/2)(r_i^L Q_i^{L,1} + X^L Q^{L,M}),\ X^L = \frac{Q_i^1 + (r_i^L - y_i^{L,1})z^L}{z^L}$
$\mathbf{C_L}$	$d_i^{L,C} = (k^j \cdot N_i^L/2)(r_i^L Q_i^{L,1} + r_{i-1} Q_i^{L,2} + Z^L Q_i^{L,M}),\ Z^L = \frac{Q_i^{L,2} + (r_i^L - y_i^{L,2})z^L}{z^L}$
$\mathbf{D_L}$	$d_i^{L,D} = (k^j \cdot N_i/2)(r_i^L \hat{Q}_i^{L,1} + (1 - r_{i-1})\hat{Q}_i^{L,2} + Z^D Q_i^{L,M}),\ Z^D = \frac{\hat{Q}_i^{L,2} + (r_i^L - \hat{y}_i^{L,2})z^L}{z^L}$

Note: **Pattern $\mathbf{A_L}$** = the left-turn queue consists of solely the upstream through vehicles.
Pattern $\mathbf{B_L}$ = discharging shockwave catches up the tail of left-turn queues under Wave 5.
Pattern $\mathbf{C_L}$ = discharging shockwave catches up the tail of left-turn queue under Wave 6; and
Pattern $\mathbf{D_L}$ = no local progression is provided between the upstream through and downstream left-turn movements.

Note that the through movement's delay during the red phase on link i should be determined by one of the three arrival patterns, as shown in Equation 19.

$$d_i = \begin{cases} d_i^A & \text{if } \delta_i^A = 1 \\ d_i^B & \text{if } \delta_i^B = 1 \\ d_i^C & \text{if } \delta_i^C = 1 \end{cases} \tag{19}$$

where δ_i^A, δ_i^B, and δ_i^C are binary variables denoting the occurrence of arrival patterns A, B, or C, respectively, depending on the green splits, offsets, and phase sequences between two adjacent intersections. These scenarios, as shown with Equation 20, are mutually exclusive, and only one pattern holds for each link.

$$\delta_i^A + \delta_i^B + \delta_i^C = 1 \tag{20}$$

Queue Evolution Process from Left-turn Vehicles

As with the through delays, one can also estimate the left-turn delays with its queue evolution patterns, which may comprise three waves. As shown in Figure 5*a*, Wave 4 and Wave 6, respectively, denote the queue formation shockwaves caused by the upstream intersection's discharged vehicles which intended a left turn at the subject intersection but were stopped by the red phase. Wave 5 refers to those queues contributed by vehicles from the upstream intersection's side street (i.e., during the red phase for through vehicles). Depending on the type of signal coordination design, one or more of the following queue evolution patterns may prevail over the arterial.

Note that considering the relatively short phase duration for the left-turn movement and the absence of local progression between left-turn and upstream through movements, left-turn queue evolution patterns may follow Pattern DL, different from the patterns for through movement. Under such a pattern, Wave 5 is most likely to be the first contributor to the left-turn queues, followed by Wave 6 constituted by through vehicles from the upstream intersection, followed by Wave 5 again.

Grounded in the logic of Equation 14 but accounting for the scenario of no left-turn progression (i.e., pattern D_L), one can formulate the left-turn queue lengths under those four patterns with Equations 21–23:

$$Q_i^L(t) = \begin{cases} Q_i^{L,0}(t) & \text{if } \delta_i^{L,D} = 0 \\ Q_i^{L,D}(t) & \text{if } \delta_i^{L,D} = 1 \end{cases} \tag{21}$$

$$Q_i^{L,0}(t) = \begin{cases} w_i^4 t & 0 \le t \le \min(y_i^{L,1}, \sigma_i^L) \\ \min[Q_i^{L,1} + w_i^5 \cdot (t - y_i^{L,1}), Q^{L,M}] & \min(y_i^{L,1}, \sigma_i^L) \le t \le \min(y_i^{L,2}, \sigma_i^L) \\ \min[Q_i^{L,2} + w_i^6 \cdot (t - y_i^{L,2}), Q^{L,M}] & \min(y_i^{L,2}, \sigma_i^L) \le t \le \sigma_i^L \end{cases} \tag{22}$$

$$Q_i^{L,D}(t) = \begin{cases} w_i^5 t & 0 \le t \le \hat{y}_i^{L,1} \\ \hat{Q}_i^{L,1} + w_i^4 \cdot (t - \hat{y}_i^{L,1}) & \hat{y}_i^{L,1} \le t \le \hat{y}_i^{L,2} \\ \hat{Q}_i^{L,2} + w_i^5 \cdot (t - \hat{y}_i^{L,2}) & \hat{y}_i^{L,2} \le t \le \sigma_i^L \end{cases} \tag{23}$$

where δ_i^D is a binary variable indicating whether or not left-turn progression exists between intersection i and the upstream through movement (equals one if without progression); $Q_i^{L,0}(t)$ and $Q_i^{L,D}$ denote the queue length which varies with the incurred patterns (i.e., pattern A_L–C_L or D_L); $\hat{y}_i^{L,1}$ is the time point when the first vehicle from the upstream through phase encounters the through queue at the location $\hat{Q}_i^{L,1}$ for queue pattern D_L; $\hat{y}_i^{L,2}$ is the time point when the last vehicle from the upstream through movement joins the through queue at the location $\hat{Q}_i^{L,2}$ under queue pattern D_L.

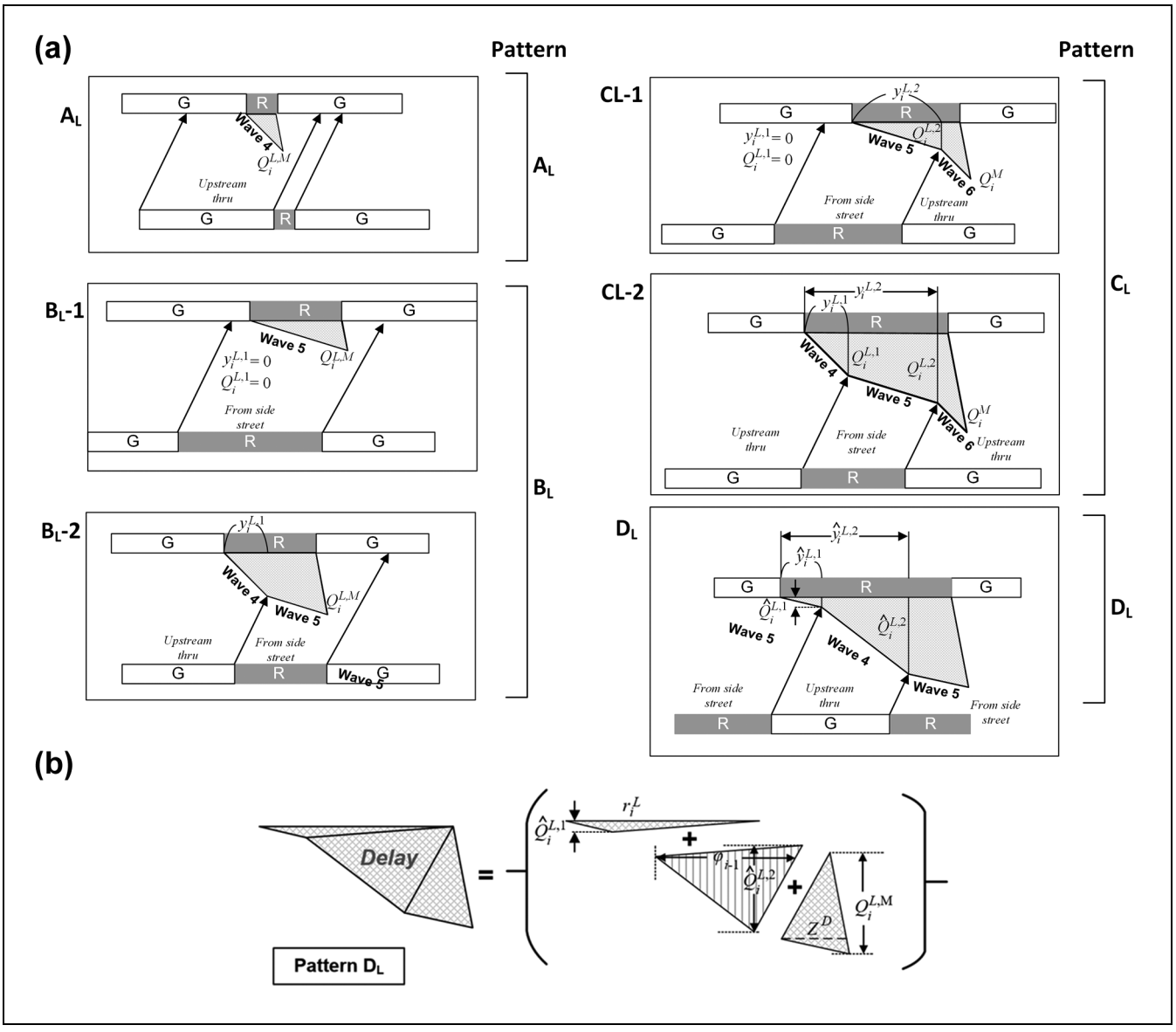

Figure 5. Arrival pattern, queue evolution, and delay computation for left-turn movements: (*a*) possible arrival patterns and queue evolutions based on the signal offsets between neighboring intersections and (*b*) delays under Pattern DL.

Delay Estimation for Left-turn Vehicles

With the evolution information of left-turn queues, the delays of left-turn vehicles can be estimated by computing the waiting time in the queues for all left-turn vehicles. For estimating delays under queue patterns A_L, B_L, and C_L, one shall follow the logic for the through movements, as shown in Table 2; variables with superscript L denote those for the left-turn movements.

As for the queue pattern D_L, mainly in the left-turn movements because of its short green duration and the lack of progression, one can compute the resulting delay, shown in Figure 5*b*, with the following expression:

$$d_i^{L,D} = \left(k^j \cdot N_i/2\right)\left(r_i^L \hat{Q}_i^{L,1} + (1 - r_{i-1})\hat{Q}_i^{L,2} + Z^D Q_i^{L,M}\right) \tag{24}$$

$$Z^D = \frac{\hat{Q}_i^{L,2} + (r_i^L - \hat{y}_i^{L,2})z^L}{z^L} \tag{25}$$

The estimated left-turn delays, shown in Equation 26, should be determined by the prevailing arrival pattern, listed in Table 2.

$$d_i^L = \begin{cases} d_i^{L,A} & \text{if } \delta_i^{L,A} = 1 \\ d_i^{L,B} & \text{if } \delta_i^{L,B} = 1 \\ d_i^{L,C} & \text{if } \delta_i^{L,C} = 1 \\ d_i^{L,D} & \text{if } \delta_i^{L,D} = 1 \end{cases} \tag{26}$$

where $\delta_i^{L,A}$, $\delta_i^{L,B}$, $\delta_i^{L,C}$ and $\delta_i^{L,D}$ are binary variables denoting the occurrence of queue patterns A_L, B_L, C_L, or D_L, between two adjacent intersections, with Equation 27 showing their mutually exclusive relations.

$$\delta_i^{L,A} + \delta_i^{L,B} + \delta_i^{L,C} + \delta_i^{L,D} = 1 \tag{27}$$

The total delays in the objective functions (Equations 1 and 4), resulting from the queues during the red phases, can be computed by summing link delays for through and left-turn movements as shown in Equation 28.

$$D^R = \sum_i (d_i + d_i^L) \tag{28}$$

Identifying Left-turn Bay Spillback and Computing the Resulting Excessive Delays

To address the impact from left-turn spillback on through vehicles, one shall first formulate the spatial and temporal relations between the left-turn queues and the turning bay. Notably, the overflows from a turning bay are most likely to incur significant delay when the spillback causes partial lane blockage to the through traffic and impedes the progression flows. As such, estimation of the delays caused by the overflow of turning traffic will focus on the impact duration contributed by Waves 4 and 6.

Let the impact duration by Wave 4 be denoted as π_i^4, which can be computed from the time difference between the spillback's starting and ending times, as shown in Figure 6*a*.

$$\pi_i^4 = \max\left\{0, y_i^{L,1} - \frac{L_i^B}{w_i^4}\right\} \tag{29}$$

where L_i^B is the left-turn bay length. The starting time of such impact is estimated with the last term, (L_i^B/w_i^4), which shows the time duration for the backward-moving queue formation shockwave at the speed w_i^4 to reach the entry of the left-turn bay. When a spillback actually occurs, its ending time can always be specified as $y_i^{L,1}$, reflecting the time of the last arriving vehicle from the upstream through movement. Note that the last term (L_i^B/w_i^4) can be a very large number to indicate that no spillback has occurred, and the entire right-hand side of Equation 29 will then be forced to equal zero.

Likewise, the same impact duration but under Wave 6, denoted as π_i^6, can be computed with Equation 30, where its ending time (σ_i^L) specified as the time when the left-turn queues completely dissipate, as shown in Figure 6*b*.

$$\pi_i^6 = \max\left\{0, \sigma_i^L - \beta_i^6\right\} \tag{30}$$

where β_i^6 is the starting time of the spillback impact under Wave 6, which can be computed with Equation 31.

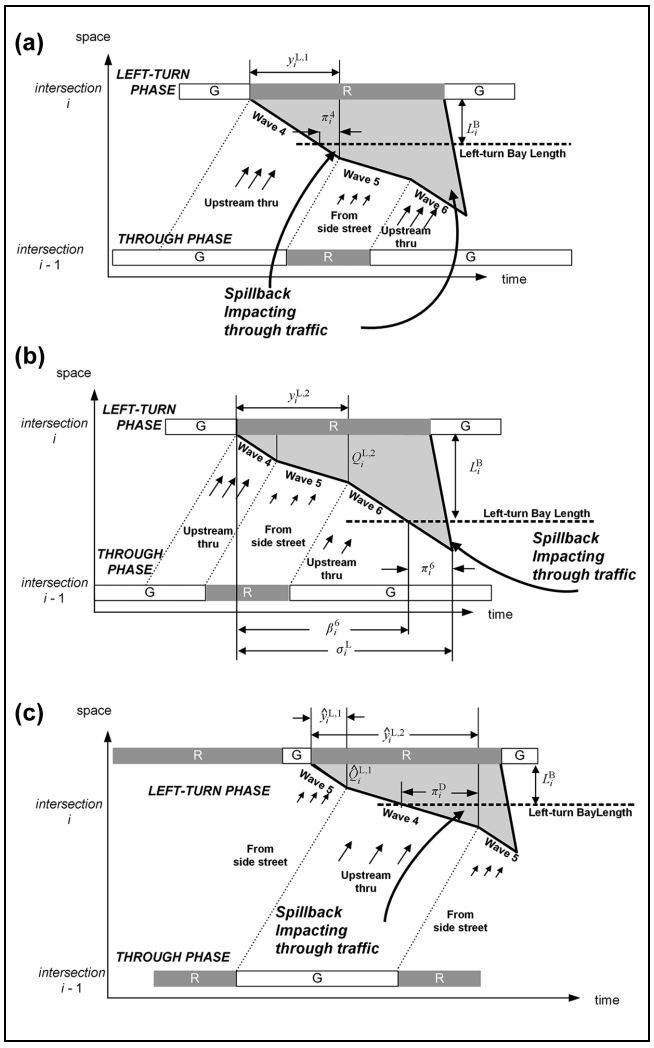

Figure 6. Impact on through traffic from left-turn spillback: (*a*) when left turn is with pattern A, B, C and Wave 4 prevails, (*b*) when left turn is with pattern C, and Wave 6 prevails, and (*c*) when left turn is with pattern D_L.

$$\beta_i^6 = \max\left(y_i^{L,2}, y_i^{L,2} + \frac{L_i^B - Q_i^{L,2}}{w_i^6}\right) \tag{31}$$

The last term of Equation 31 is for computing the required time for the left-turn queues to overflow. Starting from the first vehicle from the upstream through phase joining the tail of the left-turn queue with the length of $Q_i^{L,2}$, at the time point of $y_i^{L,2}$, the left-turn queue is accumulating at the backward propagating speed of w_i^6; the time for the tail of the left-turn queue to reach the entry of the left-turn bay can thus be computed with $\frac{L_i^B - Q_i^{L,2}}{w_i^6}$, which may yield a positive or negative value. The latter reflects that the left-turn queues have overflowed from the bay before the formation of Wave 6, so the starting time of the impact should be measured

precisely from the time point when the first vehicle from the upstream through phase arrives ($y_i^{L,2}$).

If the left-turn pattern is identified as D_L, then its impacts on through flows shall occur between the first and last upstream through vehicles' arrival times at the turning bay during its spillback period (i.e., between $\hat{y}_i^{L,1}$ and $\hat{y}_i^{L,2}$), as shown in Figure 6*c*, which can be expressed with Equation 32.

$$\pi_i^D = \begin{cases} \max\{0, \hat{y}_i^{L,2} - \max\left(\hat{y}_i^{L,1}, \hat{y}_i^{L,1} + \frac{L_i^B - \hat{Q}_i^{L,1}}{w_i^4}\right)\} & \text{if } \delta_i^{L,D} = 1 \\ 0 & \text{if } \delta_i^{L,D} = 0 \end{cases} \quad (32)$$

The last term in Equation 32, following the same logic shown in Equation 30, is specified to compute the time needed for the growing left-turn queue to reach the left-turn bay. The first through vehicle in a cycle coming from upstream to join the queue with the length of $\hat{Q}_i^{L,1}$, at timepoint of $\hat{y}_i^{L,1}$ when Wave 4 prevails, will be followed by the shockwave at the backward propagation speed w_i^4.

With the above formulations, one can compute the spillback-related delay as follows:

$$d_i^S = \pi_i^D \cdot f_{i-1}^T \cdot t^S \quad (33)$$

where f_{i-1}^T is the arrival rate of through traffic from the upstream intersection; and t^S is the additional travel time for the blocked vehicles to perform lane-changing maneuvers.

$$D^S = \sum\nolimits_i d_i^S \quad (34)$$

Note that Equations 33 and 34 produce a conservative estimate of delays from the spillback, which may trigger further traffic breakdown to degrade the quality of traffic efficiency and safety.

The Number of Approaches Experiencing Left-turn Bay Spillback

To estimate the number of approaches experiencing left-turn spillback, one can specify a binary variable, δ_i^S, which equals one if a left-turn bay spillback occurs at intersection i and zero otherwise. It is noticeable from Equation 35 that if the impact of the spillback on the through movement occurs, at least one of those terms on the right-hand side (π_i^4, π_i^6, or π_i^D) will show a positive value, forcing the left-hand side to be nonzero, and the binary variable will indicate that spillback occurs on link i; where M is a large number

$$M \cdot \delta_i^S \geq \pi_i^4 + \pi_i^6 + \pi_i^D \quad (35)$$

In brief, the number of intersection approaches experiencing queue spillback can be computed by summing up the binary indicators for all intersection approaches, as shown in Equation 36, and the results are embedded in the objective function of Model 2 (Equation 4).

$$N^S = \sum\nolimits_i \delta_i^S \quad (36)$$

The Proposed Models

With the above constraints to estimate the presence and duration of turning bay spillback, one can formulate the following models to address progression and delay concurrently.

Model 1

$$MinD^R + D^S$$

subject to

- Constraints to preserve maximized progression bandwidths, Equations 2, 3, 8–12;
- Constraints to portrait queue evolution, Equations 14–16;
- Constraints to evaluate delay on red phase, Equations 17–28 and Table 2; and
- Constraints to evaluate delay resulting from spillback, Equations 29–34.

Model 2

$$\text{Min} W^S \cdot N^S + D^R + D^S$$

s.t.

- Constraints to preserve sufficient progression bandwidths, Equations 5–12;
- Constraints to portrait queue evolution, Equations 14–16;
- Constraints to evaluate delay on red phase, Equations 17–28 and Table 2;
- Constraints to evaluate delay resulting from spillback, Equations 29–34; and
- Constraints to count the spillback occurrence of approaches, Equations 35 and 36.

Case Study

The case study, with both numerical experiments and simulation evaluations, is designed to verify the following contributions of this study:

- The proposed Model 1 (M1) with its embedded delay formulations can effectively identify the

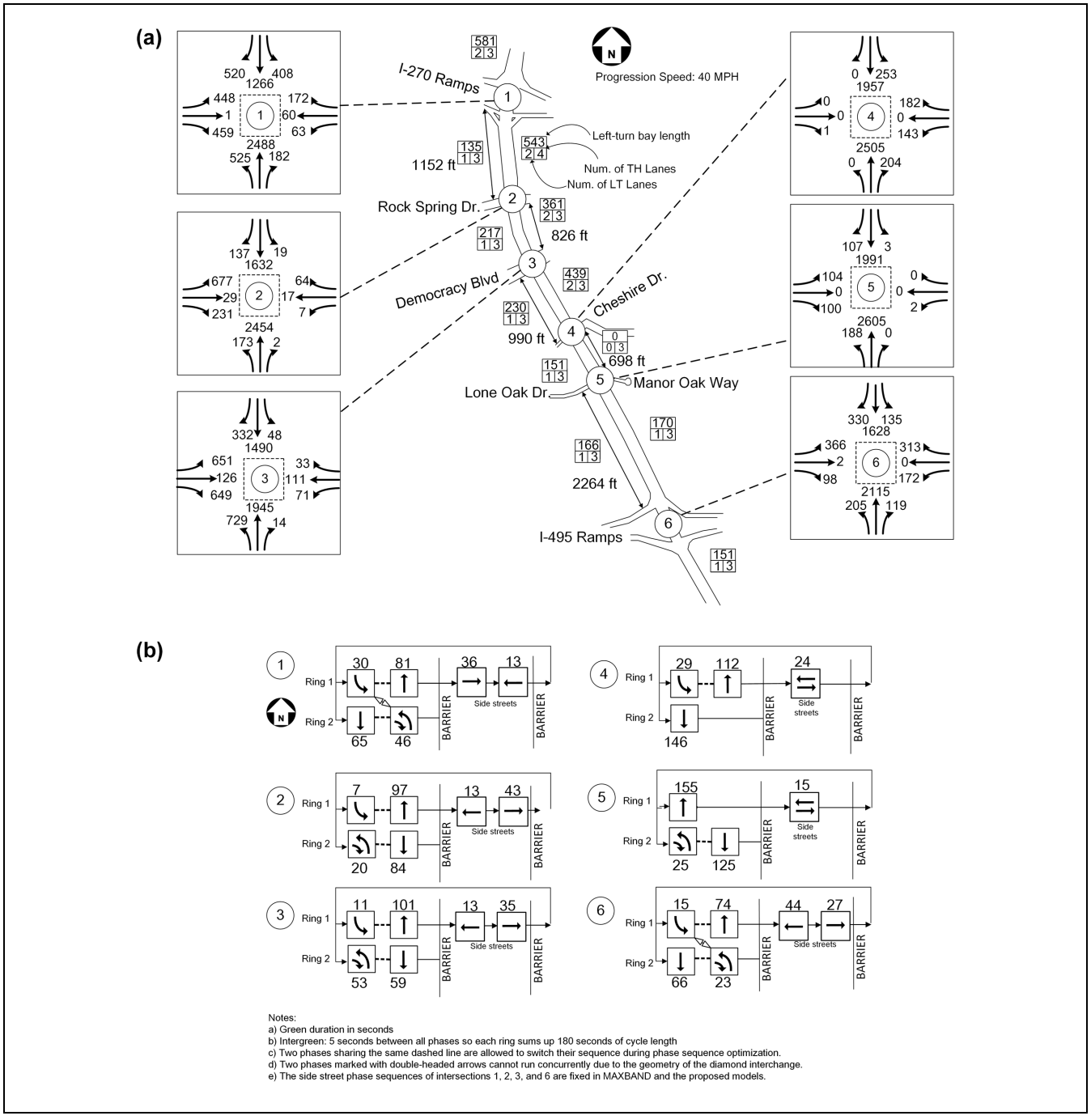

Figure 7. Key parameters of the study site: (*a*) geometric features and turning volume counts and (*b*) signal phasing plan and green splits.

progression plan with the minimal overall delay from MAXBAND's multiple optimal solutions; and

- The proposed Model 2 (M2) offers the flexibility to concurrently minimize the potential left-turn bay spillback and the overall delay by optimally reducing the maximized progression bandwidths within the pre-specified bound.

To assess the above contributions, this study has adopted the well-established TRANSYT as the benchmark and computed all resulting measures of effectiveness (MOEs) with the microscopic traffic simulation program and a real-world arterial system. Figure 7 shows the geometric features, peak hour volumes, and signal settings over an arterial of six intersections on the Old George Town Road in Bethesda, Maryland, for the case study.

Table 3. Performance Results Under the Experimental Scenarios of Different Turning Volumes

	Bandwidth				Delay				Left-turn spillback	
	Two-way (s/cycle)	SB (s/cycle)	NB (s/cycle)	Tradeoff[a] (%)	Arterial (veh-s/h)	Improv. (%)	Spillback-related (veh-s/h)	Improv. (%)	#approach	Impact duration on through movements (s/cycle)
High turning volumes										
MAXBAND	125.9*	54.2*	71.7*	na	629,163	na	8,812	na	3	54.7
M1	125.9*	54.2*	71.7*	0.0	616,604	2.0	8,771	0.5	3	54.5
M2	109.8	47.2	62.6	12.8	603,478*	4.1	5,723*	35.1	1*	21.0*
Medium turning volumes										
MAXBAND	125.9*	54.2*	71.7*	na	604,058	na	2,301*	na	2	9.5
M1	125.9*	54.2*	71.7*	0.0	583,846	3.3	2,301*	0.0	2	9.5
M2	110.3	47.2	63.1	12.4	548,968*	9.1	2,302	0.0	1*	8.3*
Low turning volumes										
MAXBAND	125.9*	54.2*	71.7*	na	527,184	na	0	0.0	0	0
M1	125.9*	54.2*	71.7*	0.0	522,878	0.8	0	0.0	0	0
M2	110.3	47.2	63.1	12.4	488,324*	7.4	0	0.0	0	0

Note: Improv. = improvement; NB = northbound; SB = southbound; s/cycle = seconds per cycle; veh-s/h = vehicle seconds per hour; NA = not applicable. Asterisk (*) represents best performances among all control strategies.
[a]Bandwidth reduction (in percent) compared with MAXBAND.

Numerical Experiments

To verify the proposed models' functions on minimizing the turning bay spillback and the overall arterial delay, the numerical experiments comprise the following three scenarios with various turning volume counts:

- High turning volume: based on the evening peak volumes, for example, the left-turn volume of 729 vehicles per hour (vph) in the northbound approach of intersection 3, as shown in Figure 7*a*.
- Medium turning volume: reducing the left-turn traffic volumes by 30 vph from the field data if exceeding 30 vph.
- Low turning volume: reducing the left-turn traffic volumes by 100 vph from the field data if exceeding 100 vph.

The commercial package Gurobi 9.1 is applied to solve the proposed models because all their embedded formulations can be converted to a standardized mixed-integer-quadratic programming system.

Table 3 shows the bandwidths, delay, and occurrence of turning bay spillbacks with the two proposed models and the benchmark progression model, MAXBAND, which does not address the delay minimization and turning bay spillback issues.

The main findings from the numerical experiment are summarized below:

Producing a Delay-Minimizing Signal Plan Among MAXBAND's Multiple Solutions

As expected, from MAXBAND's multiple solutions, the proposed M1 can indeed find the one that results in the overall minimal delay. More specifically, it is noticeable from the results in Table 3 that both models, producing the same progression bandwidths for the southbound and northbound traffic of 54.2 and 71.7 seconds per cycle (s/cycle), respectively, under medium turning volumes. Nevertheless, the proposed M1 has shown a reduction in the average delay by 2.0%, 3.3%, and 0.8%, respectively, under the experimental scenarios of high, medium, and low turning volumes, as shown in Table 3. Conceivably, the benefits of the proposed M1, featuring its capability of selecting the signal plan with minimized delay from MAXBAND's multiple solutions, will be diminishing under the traffic scenarios of low percentage of turning volumes. The results of simulation analysis also confirm that MAXBAND and M1 indeed produce similar MOEs (e.g., number of stops of 2.40 and 2.21 stops/veh, respectively, for southbound traffic along the arterial) when the major flows of the arterial are for the through movement and turning spillback is not of concern. However, if some of the arterial's intersections suffer from high turning volumes, then the signal control plan with the sole objective of maximizing the total progression bands may not achieve the desired level of performance, because the

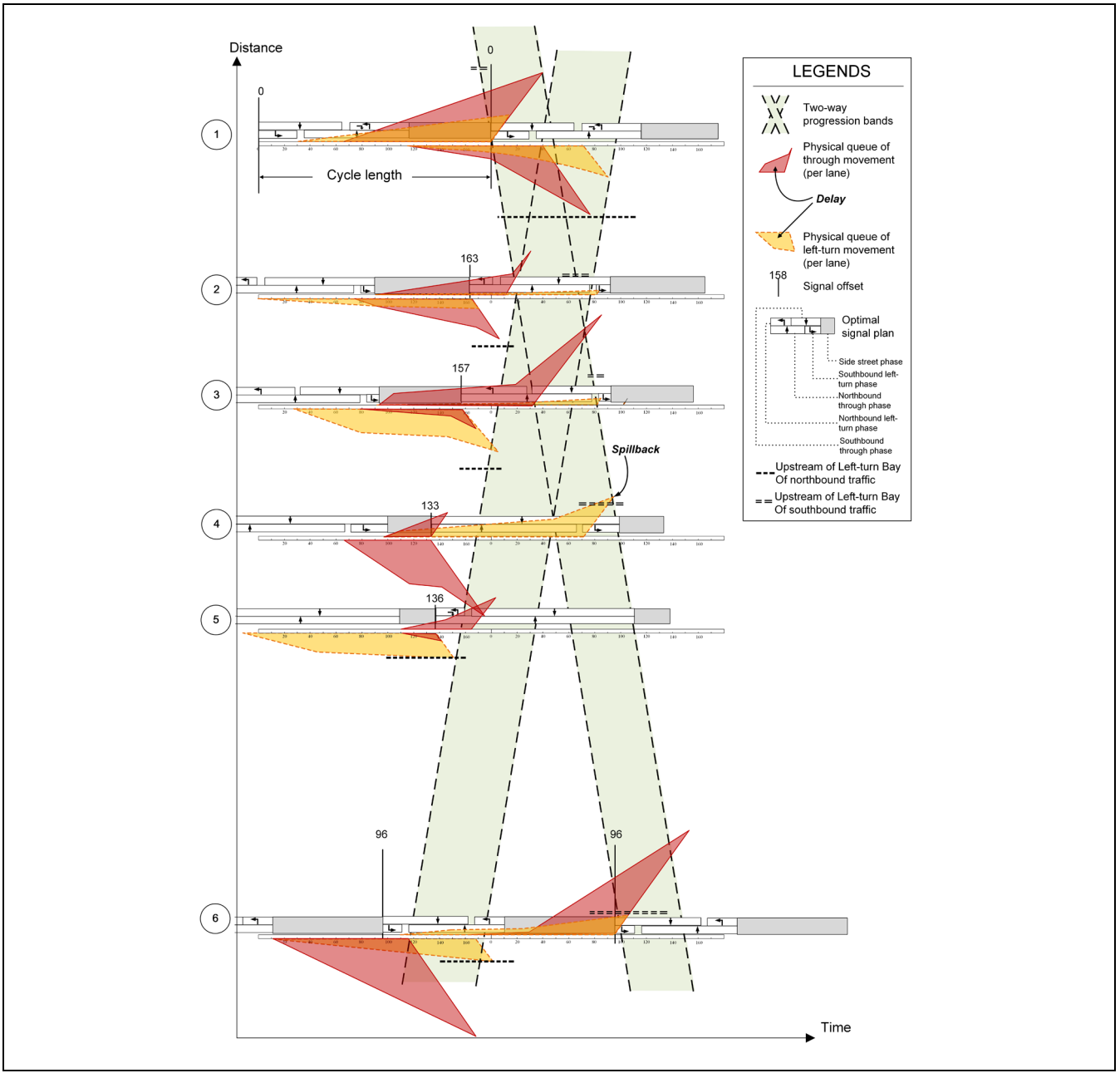

Figure 8. Time-space diagram and queue evolution with model M2 under heavy turning volumes.

turning vehicle delay and overflows are likely to impede the progression bands and degrade the overall progression effectiveness.

Circumventing Spillback and Reducing Delay

As stated previously, the M2 model is designed to enhance the methodology of progression bandwidth maximization, allowing the produced signal plan to minimize the delay and the likelihood of experiencing the turn-bay overflows by selecting the progression bandwidth, optimally reduced from their maximum level but sufficient to accommodate the longest intersection queues over the target arterial. As shown in Table 3, the delay reduction with M2 ranges from 4.1% to 9.1% under various turning volumes, compared with MAXBAND. Moreover, the signal plans by MAXBAND and M1 will result in queue spillback at three and two intersection approaches, respectively, under the traffic scenarios of heavy and medium turning volumes. In contrast, queue spillbacks only take place at only one such intersection approach under the signal progression plan by the M2

Table 4. Comparison of the Signal Plans Generated from the Four Models

	MAXBAND		Proposed M1		Proposed M2		TRANSYT16	
	Offset[a] (seconds)	LT phase sequence[b]	Offset[a] (seconds)	LT phase sequence[b]	Offset[a] (seconds)	LT phase sequence[b]	Offset[a]	LT phase sequence[b]
1	0	Lead-Lag	0	Lead-Lag	0	Lead-Lag	0	Lead-Lag
2	13	Lead-Lag	13	Lead-Lag	163	Lag-Lead	24	Lag-Lead
3	167	Lag-Lead	167	Lag-Lead	157	Lag-Lead	19	Lag-Lead
4	162	Lagging	150	Lagging	133	Lagging	3	Lagging
5	156	Leading	153	Leading	136	Leading	53	Leading
6	103	Lead-Lag	101	Lead-Lag	96	Lead-Lag	154	Lead-Lag

Note: LT = left-turn.

[a]Offsets are in the unit of seconds and are measured relative to the offset of intersection 1.

[b]Leading = leading left-turn phase; Lagging = lagging left-turn phase; Lead-Lag = leading southbound left-turn phase and lagging northbound left-turn phase; Lag-Lead = lagging southbound left-turn phase and leading northbound left-turn phase.

model, confirming that its compromised progression plan can indeed minimize the number of intersections experiencing high turning volume to suffer from the queue overflows.

Figure 8 shows the time-space diagram and time-dependent queue evolutions of all movements with M2 under heavy turning volumes. It can be observed that the southbound approach of intersection 4 has a left-turn queue longer than the turning bay and blocks the through traffic for 21.0 s/cycle (see Table 3). Notably, other intersection approaches prone to spillback (e.g., northbound approaches of intersections 3, 5, and 6) have their time-dependent left-turn queues controlled within the bay under M2.

To keep the progression bandwidth sufficiently wide to accommodate the longest intersection queue and through volume in the target arterial, the signal plan by M2, depending on the turning volumes, may not be able to relieve all turning bays from overflows. Spillback may become inevitable at some intersections where the relationship between the turning volume and bay length needs to be reconstructed properly with geometric improvements.

Reduction in Delay and Overflow Locations Under the Progression Plan with a Reduced Bandwidth

Note that the benefits of reducing total delays (e.g., from 629,163 to 603,478 veh-s/h, or 4.1%, under heavy turning volumes) and potential overflow approaches (from 3 to 1) with the M2 model are as expected at the cost of the progression bandwidth, about 12% reduction in each way. However, such a tradeoff, in practice, can better ensure the progression quality and minimize the potential impedance caused by the turning overflows to the through traffic. Some additional observations of the M2 model's benefits are summarized below:

- As shown in Table 3, the reduced bi-directional bandwidths of 16.1 s/cycle, from 125.9 to 109.8 s/cycle under heavy turning volumes, is outweighed by the benefits of decreasing the excessive delay to the through traffic caused by the turning bay queue spillback. Compared with the results under MAXBAND and M1, the affected durations with M2 have been reduced to 33.7 and 33.5 s/cycle, respectively, from 54.7 and 54.5 s/cycle to 21.0 s/cycle.
- The impacts from the turning bay spillback on the through movement are highly stochastic in nature and are likely to evolve from partial interruption to full blockage of the progression flows under some circumstances. It would thus be desirable to minimize such unexpected scenarios in practice with the proactive alternative offered by the M2 model to best ensure the progression's effectiveness.

Simulation Evaluations

To further evaluate the effectiveness of the proposed models' performance with the benchmark of TRANSYT 16, a state-of-the-art delay minimization model, this study has employed VISSIM to conduct simulation experiments, and to estimate the MOEs of signal progression plans under stochastic demand patterns and driving behaviors.

Table 4 shows the signal plans generated with the four models, and Table 5 summarizes their respective MOEs, averaged from five replications of the simulation results.

The main findings from the simulation evaluations are summarized below:

- MAXBAND, as expected, outperforms all other three models when considering the MOEs solely for the arterial's through flows. As shown in the

Table 5. Measures of effectiveness from Simulation Experiments

	MAXBAND	Proposed M1	Proposed M2	TRANSYT16
I Network performances				
Delay (s/vehicle)	124.6	122.1	117.2	121.7
(change %)[a]	na	(−2.0)	(−6.0)	(−2.4)
Number of stops (stops/vehicle)	3.00	2.93	2.85	3.08
(change %)[a]	na	(−2.3)	(−5.0)	(+2.7)
Average speed (mph)	17.42	17.63	18.04	17.67
(change %)[a]	na	(+1.2)	(+3.6)	(+1.4)
II Through traffic along arterial[b]				
Delay (s/vehicle)				
Two-way	141.2	157.2	141.8	166.1
SB (1SBTH-6SBTH)	107.3	99.6	105.5	127.9
NB (6NBTH-1NBTH)	161.9	194.5	165.0	190.0
Number of stops (stops/vehicle)				
Two-way	3.74	3.93	3.69	4.48
SB (1SBTH-6SBTH)	2.40	2.21	2.31	3.69
NB (6NBTH-1NBTH)	4.55	5.05	4.57	4.98
III Turning-out traffic from arterial[c]				
Delay(s/vehicle)				
6NBTH-3NBLT	201.5	191.3	183.8	191.4
6NBTH-2NBLT	202.8	187.9	212.8	244.8
6NBTH-1NBLT	241.3	231.0	194.1	221.3
Number of stops (stops/vehicle)				
6NBTH-3NBLT	4.74	4.46	4.19	4.79
6NBTH-2NBLT	5.11	5.03	4.26	5.25
6NBTH-1NBLT	6.24	6.01	5.42	5.92

Note: EB = eastbound; NB = northbound; LT = left-turn; TH = through traffic; na = not applicable.

[a]Compared with MAXBAND.

[b]Collected from those vehicles traversing through the entire arterial.

[c]Collected from those vehicles traversing the arterial until turning out at a particular intersection.

second set of results in Table 5, the delay of 141.2 s/veh for through traffic along the arterial with MAXBAND is less than the delays (from 141.8 to 166.1 s/veh) with other models. However, MAXBAND, which does not have the functions to minimize the total delay and to account for turning flows, yields the undesirable networkwide delay of 124.6 s/veh, the highest among all four models.

- In comparison with MAXBAND, the proposed M1, with its capability of identifying the signal progression plan with the minimum delay, can reduce its networkwide delays and number of stops, respectively, from 124.6 to 122.1 s/veh and from 3.00 to 2.93 stops/veh. Such improvements are made by shifting some offsets from MAXBAND, as shown in their offset differences at intersections 4, 5, and 6 for MAXBAND (162, 156, and 103 s, respectively) and M1 (150, 153, and 101 s, respectively) in Table 4, whereas the offsets at intersections 1 to 3 are identical (0, 13, and 167 s, respectively).
- The second proposed model, M2, effectively reduces the network delay by 6.0%, from 124.6 to 117.2 s/veh, compared with MAXBAND, as shown in the first set of results in Table 5. In addition, it also reduces the number of stops by 5.0%, from 3.00 to 2.85 stops/veh. As a result, the average network travel speed increases by 3.6%. Such improvements are consistent with the results of delay reduction revealed in the numerical experiments (see Table 3). Key contributors to such improvements include the better-coordinated turning-out flows and the minimization of queue impedance to the progression bands for the through traffic. More specifically, the signal plan under M2 has reduced the number of stops for all three major turning-out flows, for example, from 5.11 to 4.26 veh/s for vehicles turning left at intersection 2 after traveling along the entire arterial from the south boundary (6NBTH-2NBLT).
- Furthermore, by reducing the bandwidths to the level sufficient to discharge the longest queue among intersections on the arterial, vehicles under the signal plan by the M2 model can travel smoothly along the entire arterial with minimal impacts from any queue overflows. As evidenced from the second set of results in Table 5, the average delay for through vehicles over the arterial has only marginally increased from 141.2 to 141.8

s/veh, compared with MAXBAND, but the average number of stops has reduced to 3.69 from 3.74 stops/veh.
- Also, note that the average delays under the M2-generated signal plans are comparable with TRANSYT in all traffic scenarios. Specifically, the proposed M2 model has generated the lowest network average delay of 117.2 s/veh, compared with 121.7 s/veh with TRANSYT, the state-of-the-art signal model designed for networkwide delay minimization (or the performance index). The proposed M2 model, by preserving the progression bandwidths and reducing the spillback, provides smoother arterial traffic than with TRANSYT, as evidenced by the lower number of stops (3.69 and 4.48 stops/veh, respectively), the higher networkwide traffic speed (18.04 and 17.67 mph, respectively) and less average delay for the through traffic (141.8 versus 166.1 s/veh).

In summary, the proposed models can indeed reduce delay and circumvent left-turn bay spillbacks while maintaining a sufficient progression bandwidth, as demonstrated in the numerical experiments. The proposed M1 can preserve bandwidths entirely and serves to generate the signal plan with minimized delay among MAXBAND's multiple optimal solutions, while M2 offers the alternative to minimize the arterial delay and potential intersection spillback by optimally reducing the bandwidth to the level that suffices to accommodate the longest initial queue on the arterial. Simulation evaluations from VISSIM further support the findings from the numerical analyses with the following results: (i) MAXBAND can indeed produce better MOEs among all models for through vehicles traversing through the entire arterial; (ii) proposed M1 can further reduce the overall delay by minimizing the delay of turning-in and turning-out vehicles; and (iii) by minimizing the likelihood of lane blockage by turning bay overflows, the proposed M2 can generate the signal plan for the arterial that has the sufficient bandwidth to progress all through vehicles, the minimized overall delay for all arterial traffic, and the least probability to incur turning bay overflows.

Conclusion

Grounded in the core logic of MAXBAND, this study has presented two enhanced signal models to produce the optimized offsets and phase sequences for maximizing an arterial's traffic progression. The first aims to produce the progression plan that can result in the lowest overall delay for the arterial with the same total progression bandwidth, which can achieve the best efficiency for the arterial accommodating mainly through traffic flows. Aiming to minimize the delay for all traffic movements concurrently with the same control objective of progression maximization, the second model offers the flexibility to trade the bandwidth with the delay reduction, by concurrently minimizing potential left-turn overflows at the arterial's major intersections. The signal progression plan produced from the second model is especially effective for arterials comprising multiple major intersections where their turning volumes are sufficiently high and likely to incur turning bay spillback and cause excessive delay to both the through and turning flows.

The results from the numerical experiments have demonstrated that the proposed models can function as expected in minimizing the turning bay spillback and reducing the overall arterial delay under various volume scenarios. M1 is capable of producing the minimal total arterial delay under the constraint of maintaining the same maximized bandwidths as with MAXBAND. The results of M2 confirm that the arterial's total delay can be reduced by 4.1% to 9.1% under various levels of turning volumes, attributed to its functions of minimizing both the probability of left-turn bay spillback and the resulting negative impacts at the cost of a marginal reduction in the bandwidth for the through traffic. The evaluation results with simulation further confirm the effectiveness of M2 with respect to delay reduction, which produced a total delay less than that under both MAXBAND and a state-of-the-art delay minimization model, TRANSYT 16.

Under the scenario of low turning volumes as shown in the case study, note that the improvements in delay reduction and spillback prevention with the proposed models will diminish, suggesting that the signal progression plan with MAXBAND is adequate for arterials experiencing such traffic patterns. In contrast, the proposed models should be adopted in scenarios where the delays and resulting queues from the turning traffic volumes at some intersections are so significant that they need to be concurrently addressed in design of the signal progression plan when considering multiple MOEs.

In summary, the proposed models feature their unique functions to concurrently address the two conventionally competing control objectives of progression maximization and delay minimization, offering traffic engineers a potentially effective tool for the design of signal plans for major arterials, including congested intersections and heavy turning flows. The proposed models are also expected to serve as the backbone of a real-time adaptive arterial signal system when real-time traffic conditions and hardware equipment are available.

Future extension of this research will be to apply the proposed models for their real-time operations under proactive or responsive control environments. In addition, another extension of interest is to concurrently include the green split and cycle length optimization in design of the signal progression plan that can best

prevent the occurrence of turning bay spillbacks and maximize the delay reduction.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Y.-H. Chen, Y. Cheng, G.-L. Chang; data collection: Y.-H. Chen; analysis and interpretation of results: Y.-H. Chen, Y. Cheng; draft manuscript preparation: Y. Cheng, Y.-H. Chen, G.-L. Chang. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Yao Cheng https://orcid.org/0000-0002-0513-0272

References

1. Little, J. D., M. D. Kelson, and N. M. Gartner. MAXBAND: A Program for Setting Signals on Arteries and Trangular Networks. *Transportation Research Record: Journal of the Transportation Research Board*, 1981. 795: 40–46.
2. Morgan, J. T., and J. D. C. Little. Synchronizing Traffic Signals for Maximal Bandwidth. *Operation Research, Vol.* 12, No. 6, 1964, pp. 896–912.
3. Little, J. D. C. The Synchronization of Traffic Signals by Mixed-Integer Linear Programming. *Operation Research, Vol.* 14, No. 3, 1966, pp. 568–594.
4. Gartner, N. H., S. F. Assman, F. Lasaga, and D. L. Hou. A Multi-Band Approach to Arterial Traffic Signal Optimization. *Transportation Research Part B, Vol.* 25, No. 1, 1991, pp. 55–74.
5. Chaudhary, N. A., V. G. Kovvali, C. L. Chu, J. Kim, and S. M. Alam. Software for Timing Signalized Arterials. Report FHWA/TX-03/4020-1. Texas Transportation Institute, The Texas A&M University System, College Station, TX, 2002.
6. Yang, X., Y. Cheng, and G. L. Chang. A Multi-Path Progression Model for Synchronization of Arterial Traffic Signals. *Transportation Research Part C: Emerging Technologies,* Vol. 53, 2015, pp. 93–111.
7. Arsava, T., Y. Xie, and N. H. Gartner. Arterial Progression Optimization Using OD-BAND: Case Study and Extensions. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2558: 1–10.
8. Yan, H., F. He, X. Lin, J. Yu, M. Li, and Y. Wang. Network-Level Multiband Signal Coordination Scheme Based on Vehicle Trajectory Data. *Transportation Research Part C: Emerging Technologies*, Vol. 107, 2019, pp. 266–286.
9. Gartner, N. H., and C. Stamatiadis. Arterial-Based Control of Traffic Flow in Urban Grid Networks. *Mathematical and Computer Modelling,* Vol. 35, No. 5–6, 2002, pp. 657–671.
10. Yang, X., G. L. Chang, and S. Rahwanji. Development of a Signal Optimization Model for Diverging Diamond Interchange. *Journal of Transportation Engineering,* Vol. 140, No. 5, 2014, p. 04014010.
11. Cheng, Y., G. L. Chang, and S. Rahwanji. Concurrent Optimization of Signal Progression and Crossover Spacing for Diverging Diamond Interchanges. *Journal of Transportation Engineering, Part A: Systems,* Vol. 144, No. 3, 2018, p. 04018001.
12. Yang, X., Y. Cheng, and G. L. Chang. Operational Analysis and Signal Design for Asymmetric Two-Leg Continuous-Flow Intersection. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2553: 72–81.
13. Robertson, D. I. *TRANSYT: A Traffic Network Study Tool*. TRRL Report LR 253. Road Research Laboratory, Crowthorne, Berkshire, UK, 1969.
14. Michalopoulos, P. G., G. Stephanopoulos, and G. Stephanopoulos. An Application of Shock Wave Theory to Traffic Signal Control. Transportation Research Part B: *Methodological*, Vol. 15, No. 1, 1981, pp. 35–51.
15. Lo, H. K., E. Chang, and Y. C. Chan. Dynamic Network Traffic Control. *Transportation Research Part A: Policy and Practice*, Vol. 35, No. 8, 2001, pp. 721–744.
16. Stevanovic, A., P. T. Martin, and J. Stevanovic. VisSim-Based Genetic Algorithm Optimization of Signal Timings. *Transportation Research Record: Journal of the Transportation Research Board*, 2007. 2035: 59–68.
17. Kashani, H. R., and G. N. Saridis. Intelligent Control for Urban Traffic Systems. *Automatica,* Vol. 19, 1983, pp. 191–197.
18. Yun, L., and B. Park. Application of Stochastic Optimization Method for an Urban Corridor. *Proc., Winter Simulation Conference*, Monterey, CA, 2006, pp. 1493–1499.
19. Liu, Y., and G. L. Chang. An Arterial Signal Optimization Model for Intersections Experiencing Queue Spillback and Lane Blockage. *Transportation Research Part C: Emerging Technologies*, Vol. 19, No. 1, 2011, pp. 130–144.
20. Wallace, C. E., and K. G. Courage. Arterial Progression – New Design Approach. *Transportation Research Record: Journal of the Transportation Research Board*, 1982. 881: 53–59.
21. Cohen, S. L. Concurrent Use of MAXBAND and TRANSYT Signal Timing Programs for Arterial Signal Optimization. *Transportation Research Record: Journal of the Transportation Research Board*, 1983. 906: 81–84.
22. Cohen, S. L., and C. C. Liu. The Bandwidth-Constrained TRANSYT Signal Optimization Program. *Transportation Research Record: Journal of the Transportation Research Board*, 1986. 1057: 1–7.
23. Chang, E. C. P., C. J. Messer, and B. G. Marsden. Reduced-Delay Optimization and other Enhancements in

the PASSER II-84 Program (No. HS-039 134). *Transportation Research Record: Journal of the Transportation Research Board*, 1985. 1005: 80–89.

24. Chang, E. C. P., and C. J. Messer. Minimum Delay Optimization of a Maximum Bandwidth Solution to Arterial Signal Timing. *Transportation Research Record: Journal of the Transportation Research Board*, 1985. 1005: 89–95.
25. Lan, C. J., C. Messer, N. A. Chaudhary, and E. C. P. Chang. Compromise Approach to Optimize Traffic Signal Coordination Problems During Unsaturated Conditions. *Transportation Research Record: Journal of the Transportation Research Board*, 1992. 1320: 112–160.
26. Chen, Y. H., Y. Cheng, and G. L. Chang. Concurrent Progression of Through and Turning Movements for Arterials Experiencing Heavy Turning Flows and Bay-Length Constraints. *Transportation Research Record: Journal of the Transportation Research Board*, 2019. 2673: 525–537.
27. Chen, C., X. Che, W. Huang, and K. Li. A Two-Way Progression Model for Arterial Signal Coordination Considering Side-Street Turning Traffic. *Transportmetrica B: Transport Dynamics*, Vol. 7, No. 1, 2019, pp. 1627–1650.
28. Lighthill, M. J., and G. B. Whitham. On Kinematic Waves II. A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, Vol. 229, No. 1178, 1955, pp. 317–345.
29. Richards, P. I. Shock Waves on the Highway. *Operations Research, Vol.* 4, No. 1, 1956, pp. 42–51.
30. Michalopoulos, P. G., G. Stephanopoulos, and V. B. Pisharody. Modeling of Traffic Flow at Signalized Links. *Transportation Science*, Vol. 14, No. 1, 1980, pp. 9–41.
31. Wu, X., and H. X. Liu. A Shockwave Profile Model for Traffic Flow on Congested Urban Arterials. *Transportation Research Part B: Methodological*, Vol. 45, No. 10, 2011, pp. 1768–1786.
32. Cho, H. J., M. T. Tseng, and M. C. Hwang. Using Detection of Vehicular Presence to Estimate Shockwave Speed and Upstream Traffics for a Signalized Intersection. *Applied Mathematics and Computation, Vol.* 232, 2014, pp. 1151–1165.
33. Ramezani, M., and N. Geroliminis. Queue Profile Estimation in Congested Urban Networks With Probe Data. *Computer-Aided Civil and Infrastructure Engineering*, Vol. 30, No. 6, 2014, pp. 414–432.
34. Liang, Y., Z. Wu, J. Li, F. Li, and Y. Wang. Shockwave-Based Queue Length Estimation Method for Presignals for Bus Priority. *Journal of Transportation Engineering, Part A: Systems*, Vol. 144, No. 9, 2018, p. 04018057.
35. Chen Y.H., Y. Cheng, and G.L. Chang. Incorporating bus delay minimization in design of signal progression for arterials accommodating heavy mixed-traffic flows. *Journal of Intelligent Transportation Systems*, 2021, pp. 1–30. https://doi.org/10.1080/15472450.2021.2002149.