Adaptive Signal Control at Partially Connected Intersections: A Stochastic Optimization Model for Uncertain Vehicle Arrival Rates

Shaocheng JIA [a], S.C. WONG [a,*] , and Wai WONG [b,*]

[a]*Department of Civil Engineering, The University of Hong Kong, Hong Kong, China*

[b]*Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch 8041, New Zealand*

**Co-corresponding author*

Abstract

Optimizing traffic signal control is crucial for improving efficiency in congested urban environments. Current adaptive signal control systems predominantly rely on on-road detectors, which entail significant capital and maintenance costs, thereby hindering widespread implementation. In this paper, a novel connected vehicle (CV)-based adaptive signal control (CVASC) framework is proposed that optimizes signal plans on a cycle-by-cycle basis without the need for on-road detectors, leveraging partial CV data. The framework comprises a consequential system delay (CSD) model, deterministic penetration rate control (DPRC), and stochastic penetration rate control (SPRC). The CSD model analytically estimates vehicle arrival rates and, consequently, the total junction delay, utilizing CV penetration rates as essential inputs. Employing the CSD model without considering CV penetration rate uncertainty results in fixed vehicle arrival rates and leads to DPRC. On the other hand, incorporating CV penetration rate uncertainty accounts for uncertain vehicle arrival rates, establishing SPRC, which poses a high-dimensional, non-convex, and stochastic optimization problem. An analytical stochastic delay model using generalized polynomial chaos expansion is proposed to efficiently and accurately estimate the mean, variance, and their gradients for the CSD model within SPRC. To solve DPRC and SPRC, a gradient-guided golden section search algorithm is introduced. Comprehensive numerical experiments and VISSIM simulations demonstrate the effectiveness of the CVASC framework, emphasizing the importance of accounting for CV penetration rate uncertainty and uncertain vehicle arrival rates in achieving optimal solutions for adaptive signal optimizations.

Keywords: *CV-based adaptive signal control, CV penetration rate uncertainty, stochastic optimization, generalized polynomial chaos expansion, gradient-guided golden section search algorithm*

1. Introduction

The advent of fifth-generation (5G) mobile communication technologies has facilitated the development of connected vehicle (CV) systems. These vehicles can exchange real-time traffic information (including time, speed, and location) among themselves as CV data. This presents promising opportunities for developing novel solutions to complex transportation challenges. However, as such connectivity is still in its early stage, only a fraction of vehicles on the road are equipped with CV capabilities. This partial connectivity is anticipated to persist for an extended period owing to

diverse technological, societal, and ethical factors. During this prolonged transition period toward full CV deployment, accurately estimating the CV penetration rate, defined as the ratio of CV volume to the total traffic volume, is crucial for the effective planning and implementation of CV-based transportation applications.

Links with on-road detectors can directly provide total traffic volume; however, many links within a network lack such detectors, making accurate CV penetration rate estimation challenging. Furthermore, even in the presence of detectors, intermittent maintenance or technical issues may compromise their reliability. Several approaches have been explored to estimate CV penetration rates on links without detectors. One approach entails using the average of the CV penetration rate distribution, derived from CV penetration rates sampled from nearby links with detectors, as an estimation (Wong and Wong, 2015, 2016a, 2019; Wong, Wong, and Liu, 2019). However, this method relies on the assumption of independent and identically distributed CV penetration rates. Meng et al. (2017a) developed an empirical model to explain the spatial variance by leveraging land-use properties, although the model depends on locally collected data. In contrast, a more promising approach involves estimating the CV penetration rate solely using CV data. For example, under the assumption of Poisson arrival, Comert (2016) derived multiple analytical models for estimating the CV penetration rate. To address the limitations associated with Poisson arrival assumptions, Wong et al. (2019) proposed the single-source data penetration rate estimator (SSDPRE) to unbiasedly estimate the CV penetration rate regardless of arrival patterns. Additionally, other methods based on the maximum likelihood approach and leveraging the distribution of the stopping positions of vehicles in queues have been developed for estimating CV penetration rates (Zhao et al., 2019a, 2019b, 2022). Nevertheless, these estimates are point estimators, meaning that they do not account for the uncertainty in the CV penetration rate. Utilizing point estimators for transport modeling and system optimization may result in biased models and suboptimal solutions (Yin, 2008; Wong and Wong, 2015, 2016, 2019; Wong, Wong, and Liu, 2019).

Building upon the SSDPRE, Jia et al. (2023) derived the probabilistic penetration rate (PPR) model to explicitly model the uncertainty in the CV penetration rate. This work represents a significant paradigm shift from deterministic analysis to stochastic analysis in CV penetration rate estimation. It provides a robust probabilistic analysis framework for stochastic modeling and optimization in CV-based transportation systems, offering valuable insights into various CV-related challenges. Additionally, further considering complex residual-vehicle effects, Jia et al. (2024a, 2024b) proposed a Markov-constrained queue length (MCQL) model for accurately estimating constrained queue length, which is an essential input for the PPR model. This development ensures that the PPR model can effectively handle all under-saturation scenarios regardless of the presence of residual vehicles.

The aforementioned methods for estimating the CV penetration rate have paved the way for numerous beneficial applications. For example, Comert and Cetin (2009, 2011), Comert (2013), and Hao et al. (2014) have utilized the CV penetration rate as input to develop a range of methods for estimating queue length. Recovering traffic information for non-CVs using CV signals, such as

locations, speeds, and travel times, is extremely important for optimal traffic operations and has thus been extensively investigated (Jenelius et al., 2013, 2015; Rahmani et al., 2015; Khan et al., 2017; Iqbal et al., 2018). Other relevant applications include traffic flow and density estimation (Geroliminis and Daganzo, 2008; Ambühl and Menendez, 2016; Du et al., 2016; Wong and Wong, 2016c; Wong, Wong, and Liu, 2021), origin–destination estimation (Yang et al., 2017; Wang et al., 2020; Cao et al., 2021), and traffic incident impact evaluation (Wong and Wong, 2016b).

Among various applications, CV-based adaptive signal control (CVASC) has received considerable attention over the past decade. Feng et al. (2015) proposed a phase allocation algorithm to dynamically adjust signal sequences and timings in a connected environment. This approach involves dividing the problem into two sub-problems: a dynamic programming problem for signal sequence determination and a utility minimization problem for phase duration optimization. The proposed approach requires the availability of a complete arrival table, which must be estimated using the accessible CV data. To enhance the performance of adaptive signal control in scenarios with low CV penetration rates, dedicated models utilizing only a few CV trajectories have been developed to accurately estimate traffic delay and arrival information (Feng et al., 2018; Wang et al., 2021). In addition, multi-mode adaptive traffic signal control systems that leverage both CV data and existing detector data have been explored (Rafter et al., 2020; Al Islam et al., 2020). Liang et al. (2020a) proposed a real-time traffic control scheme to balance efficiency and equity (measured by average and maximum vehicle delays, respectively) using CV data. Mo et al. (2022) investigated the application of reinforcement learning to traffic signal control based on CV data. Despite these advancements, current research in adaptive signal control based on CV data has predominantly relied on deterministic models, overlooking uncertainties in traffic state estimation. Given the nonlinear and time-varying nature of transportation systems, deterministic optimization that does not consider variability can lead to suboptimal solutions in traffic signal control (Yin, 2008; Jia et al., 2023, 2024a, 2024b). Therefore, incorporating uncertainties in traffic state estimations is essential for achieving optimal solutions in adaptive signal control within a connected environment.

The present paper proposes a novel CVASC framework that optimizes signal plans on a cycle-by-cycle basis to minimize total junction delay using only available partial CV data. To analytically estimate vehicle arrival rates, and hence total junction delay, a consequential system delay (CSD) model with CV penetration rates as essential inputs is derived. However, the estimated CV penetration rates are subject to uncertainty. Ignoring this uncertainty yields fixed vehicle arrival rates and deterministic penetration rate control (DPRC). In contrast, incorporating the uncertainty associated with the CV penetration rate accounts for uncertain vehicle arrival rates and establishes stochastic penetration rate control (SPRC). Given the high-dimensional, non-convex, and stochastic nature of SPRC, efficiently solving such an optimization problem poses significant challenges. To address this, an analytical stochastic delay (ASD) model integrating generalized polynomial chaos expansion (gPCE) (Xiu, 2010) is proposed to accurately estimate the mean and variance of total junction delay and their gradients in

SPRC. Subsequently, an efficient gradient-guided golden section search (G^3S^2) algorithm is introduced as the solution method. Extensive numerical experiments and VISSIM simulations based on a real-world intersection layout extracted from the Next Generation Simulation (NGSIM) dataset (Federal Highway Administration, 2006) demonstrate the applicability and efficiency of the proposed models and the significance of incorporating CV penetration rate uncertainty and uncertain vehicle arrival rates into adaptive signal control.

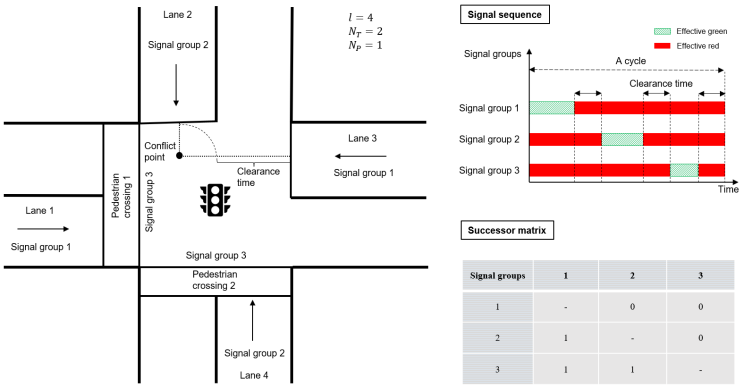
The remainder of this paper is structured as follows. Section 2 defines the problem. Section 3 systematically presents the CVASC framework. Sections 4 and 5 detail the derivations of the ASD model and the G^3S^2 algorithm, respectively. Section 6 provides a comprehensive set of realistic simulations. Finally, Section 7 concludes the paper.

2. Problem Statement

Consider a signalized isolated intersection with an arbitrary number of approaches, all without any detectors. Each approach consists of a specific number of lanes leading to the junction. The total number of approaching lanes for the intersection, denoted as l, is the sum of the lanes in all of the approaches. The numbers of traffic and pedestrian groups are N_T and N_P, respectively. Given the geometric details of the junction, let $\mathbf{\Omega} \in \mathbb{R}^{(N_T+N_P)\times(N_T+N_P)}$ and $\mathbf{T} \in \mathbb{R}^{(N_T+N_P)\times(N_T+N_P)}$ be the successor matrix and the clearance time matrix, respectively. Ω_{ij} is the element at the i^{th} row and the j^{th} column in $\mathbf{\Omega}$, where

$$\Omega_{ij} = \begin{cases} 0 & if\ group\ i\ precedes\ group\ j \\ 1 & otherwise \end{cases}. \tag{1}$$

Additionally, t_{ij} is the element at the i^{th} row and the j^{th} column in $\mathbf{T}$, and it indicates the minimum clearance time during a transition from signal group i to j. The set of incompatible signal group pairs is denoted as $\mathbf{\Psi}$. A sample intersection with two traffic groups and one pedestrian group is used to illustrate the above definitions (Figure 1). In the transition period of a mixed driving environment involving both CVs and non-CVs within the network, only CV data are available. Given the aforementioned settings and constraints, this paper aims to develop a CVASC framework that optimizes cycle-by-cycle adaptive signal control by minimizing the total junction delay solely using the available partial CV data.



Signal groups	1	2	3
1	-	0	0
2	1	-	0
3	1	1	-

Figure 1. Illustration of definitions using a sample intersection.

The total junction delay, denoted as D, is the sum of traffic delays across all lanes. At the end of cycle k, the total junction delay for cycle $k+1$ has to be estimated. The vehicle arrival rate in lane m in cycle $k+1$, denoted as $q_{m,k+1}$, is the essential input for estimating the traffic delay in lane m in cycle $k+1, \forall m \in [1, l]$. With the assumption of similar traffic demands between consecutive cycles and an identical average arrival rate in lane m in cycle $k+1$ to that in cycle k, $q_{m,k+1}$ can be estimated as follows:

$$q_{m,k+1} = \frac{n_{m,k}}{C_k} + \bar{q}_{m,k}\left(1 - p_{m,k}\right), \forall m \in [1, l], \tag{2}$$

where $n_{m,k}$, C_k, $\bar{q}_{m,k}$, and $p_{m,k}$ represent the number of CVs in lane m in cycle k, length of cycle k, average arrival rate in lane m estimated at the end of cycle k, and CV penetration rate in lane m in cycle k, respectively (Jia et al., 2023, 2024a, 2024b). While $n_{m,k}$, C_k, and $\bar{q}_{m,k}$ are constants, the estimated $p_{m,k}$ could be subject to uncertainty. The CV penetration rate is defined as the ratio of the number of CVs in a cycle to the total number of vehicles in that cycle. Wong et al. (2019) proposed a novel SSDPRE method to unbiasedly estimate the CV penetration rate based on the number of CVs and the total number of vehicles in a constrained queue deduced from the stopping location of the last CV. The variance of the estimator quantifies the uncertainty in CV penetration rate and thus serves as an indicator of its variation range. CV penetration rate uncertainty can generally be attributed to one of the following four key factors: (1) variations in the total number of vehicles, (2) variations in the number of CVs, (3) permutations of CVs and non-CVs, and (4) the presence of residual vehicles from previous cycles. Jia et al. (2023) analytically quantified CV penetration rate uncertainty due to (1) variations in the total number of vehicles, (2) variations in the number of CVs, and (3) permutations of CVs and non-CVs. Subsequently, Jia et al. (2024a, 2024b) extended their analysis by further quantifying CV penetration rate uncertainty due to (4) the residual vehicles from the previous cycles. In general, CV penetration rate uncertainty decreases as volume-to-capacity (V/C) ratio increases, due to the richer information available from partially observed constrained queues. Additionally, when the underlying CV penetration rate is either very low or very high, the uncertainty approaches zero due to the near absence or abundance of observed CVs, respectively. As the total junction delay D is a function of $q_{m,k+1}$, which itself depends on $p_{m,k}$, D is inherently subject to uncertainty associated with $p_{m,k}$. Ignoring other constants, the abstract form of D, dependent on a set of group-based control variables, can be written as

$$D = f\left(\theta_1, \dots, \theta_i, \dots, \theta_{N_T+N_P}, \phi_1, \dots, \phi_i, \dots, \phi_{N_T+N_P}, \zeta \middle| p_{1,k}, \dots, p_{m,k}, \dots, p_{l,k}\right), \tag{3}$$

where $\theta_i, \forall i \in [1, N_T + N_P]$ represents the time from the cycle origin to the start of an actual green signal for control group i divided by the cycle time, $\phi_i, \forall i \in [1, N_T + N_P]$ represents the duration of the actual green signal for control group i divided by the cycle time, and ζ is the reciprocal of the cycle length (Wong, 1996; Silcock, 1997; Lee et al., 2017a, 2017b).

It is important to emphasize that the total junction delay, D, largely depends on vehicle arrival rates, which in turn are influenced by the CV penetration rates. Evaluating D and optimizing signal plans are straightforward if $p_{m,k}, \forall m \in [1, l]$, is taken as a point estimator without considering its variability, $Var(p_{m,k})$. In this case, the estimated vehicle arrival rates are fixed and Eq. (3) simplifies into a deterministic model. However, deterministic models can be biased and may lead to suboptimal solutions in system optimizations (Yin, 2008; Wong and Wong, 2015, 2016a, 2019; Jia et al., 2023, 2024a, 2024b). In a simulation study of a simple two-approach intersection, Jia et al. (2023) demonstrated that incorporating the uncertainty of the CV penetration rate into a CV-based adaptive signal optimization problem could result in significant improvements. Specifically, a 15.1% reduction in average delay, a 15.3% decrease in maximum delay, and a substantial 45.5% drop in delay variance were observed. Considering both the mean and uncertainty of $p_{m,k}$ turns it into a random variable, accounting for uncertain vehicle arrival rates and making Eq. (3) a stochastic model. With numerous control variables and traffic lanes for a complex intersection, efficiently evaluating $E(D)$ and $Var(D)$ for cycle-by-cycle adaptive signal control becomes challenging. Furthermore, obtaining the partial derivatives of $E(D)$ and $Var(D)$ for developing gradient-based optimization algorithms increases the complexity. While Monte Carlo sampling (MCS) is convenient for diverse evaluations, its low computational efficiency hinders real-time applications. Thus, the key research problem of this paper is to develop a CVASC framework with models that can efficiently estimate $E(D)$, $Var(D)$, and their partial derivatives to achieve cycle-by-cycle adaptive signal control solely using CV data. In addition to the terminology and notation defined above, a glossary and a table of symbols are provided in **Appendix A**.

3. CV-based Adaptive Signal Control (CVASC) Framework

This section presents the formulation of the CV-based adaptive signal control (CVASC) framework, starting with the consequential system delay (CSD) model that estimates the total junction delay. A set of constraints is introduced for signal optimization. Lastly, the deterministic penetration rate control (DPRC) and stochastic penetration rate control (SPRC) models are developed.

Developing the CVASC framework under uncertain CV penetration rates requires estimating several critical quantities, including the uncertainty of the CV penetration rate in lane m in cycle k, $Var(p_{m,k})$, the average arrival rate in lane m in cycle k, $\bar{q}_{m,k}$, and the number of holding vehicles in lane m in cycle k, $R_{m,k}$, where holding vehicles represent vehicles that, based on their projected trajectories using cruise speed, should have been discharged but are instead held by the system and remain undischarged at the end of a cycle (Jia et al., 2024c). Previous studies by Wong et al. (2019) and Jia et al. (2023, 2024a, 2024b, 2024c) have demonstrated that these quantities can be estimated in situations with partial CV trajectories. Thus, they are considered known inputs for the proposed

framework in this study. **Appendix B** provides concise overviews of the estimation methods for these essential quantities.

In addition, Figure 2 provides a step-by-step flowchart summarizing the implementation procedure of the proposed CVASC framework. In Step 1, CV trajectories are extracted from real-time traffic data. In Step 2, these CV trajectories are used to estimate essential traffic parameters. In Step 3, the delay estimation model is established (Section 3.1). In Steps 4 and 5, the optimization constraints are determined and the control schemes are defined, respectively (Sections 3.2–3.4). Finally, in Step 6, the signal optimization problem is solved (Sections 4 and 5), and the optimized signal plan is implemented in the traffic system. In practice, this process is continuously repeated throughout the control period.

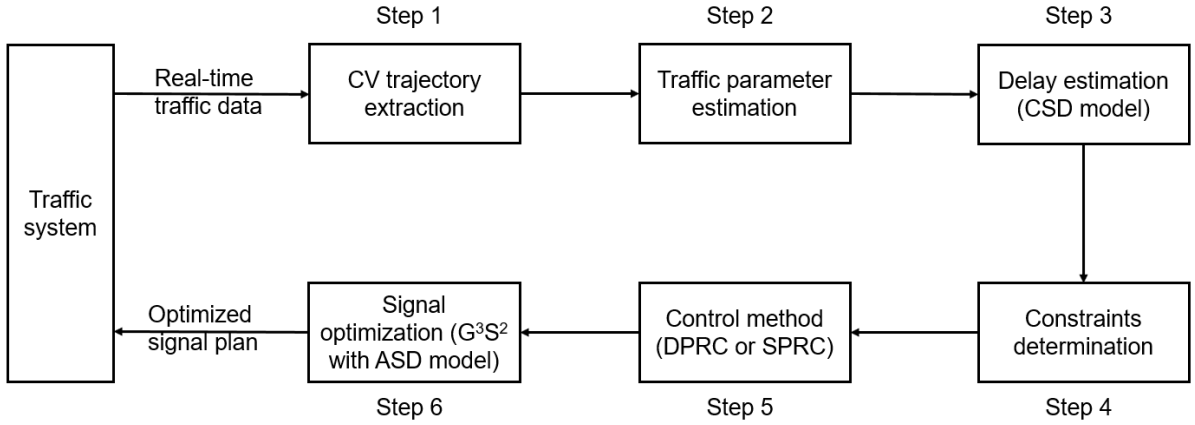


Figure 2. Flowchart illustrating the process of CVASC framework implementation.

3.1. Consequential system delay model

The CSD model estimates both the traffic delay in cycle $k+1$ and the consequential delay induced by the holding vehicles at the end of cycle $k+1$, as presented in **Propositions 1** and **2**, respectively. The estimated delay is determined by the area enclosed between the cumulative arrival curve and the cumulative departure curve. Typically, the cumulative departure curve is closely modeled by a straight line with a slope equal to the saturation flow rate, representing the maximum discharge rate of queued vehicles. The cumulative arrival curve, on the other hand, depends on the vehicle arrival rate and the specific vehicle arrival pattern. The estimated delay is mainly governed by the vehicle arrival rate, which dictates the overall slope of the cumulative arrival curve and largely determines the total number of arriving vehicles in a cycle. As illustrated in Eq. (2), the estimated vehicle arrival rate depends on the CV penetration rate. Ignoring the CV penetration rate uncertainty results in fixed vehicle arrival rates and DPRC. Considering the crucial role of the vehicle arrival rate in estimating delay, incorporating CV penetration rate uncertainty accounts for the uncertainty in vehicle arrival rates and, consequently, the uncertainty in the overall slope of the cumulative arrival curve, leading to the establishment of SPRC. In comparison, the impact of the specific vehicle arrival pattern on the estimated delay is relatively minor compared to the vehicle arrival rate, as the total number of arriving vehicles is

already largely governed by the vehicle arrival rate. In the absence of the detailed vehicle arrival pattern, particularly for non-CVs, assuming a uniform vehicle arrival pattern serves as a reasonable and accurate first-order approximation. **Proposition 1** provides the traffic delay for any lane m in cycle $k+1$, denoted as $d_{m,k+1}$.

Proposition 1. Under the assumption of a uniform vehicle arrival pattern, the traffic delay for any lane m in cycle $k+1$, denoted as $d_{m,k+1}$, is estimated as $d_{m,k+1}^{(1)}$ or $d_{m,k+1}^{(2)}$ depending on whether the lane is controlled by a signal group ending with effective red or green, respectively. $d_{m,k+1}^{(1)}$ and $d_{m,k+1}^{(2)}$ under different conditions are detailed in Tables 1 and 2, and the parameters C, r , r_1, r_2, g, g_1, and g_2 in Eqs. (4)–(10) respectively represent the cycle length, total effective red, effective red before effective green, effective red after effective green, total effective green, effective green before effective red, and effective green after effective red for the associated signal group controlling lane m.

Table 1. Analytical formulae for $d_{m,k+1}^{(1)}$.

Conditions	$\boldsymbol{d_{m,k+1}^{(1)}}$
$q_{m,k+1} \le \frac{sg - R_{m,k}}{r_1 + g}$	$\frac{1}{2}(2R_{m,k} + q_{m,k+1}r_1)r_1 + \frac{(R_{m,k} + q_{m,k+1}r_1)^2}{2(s - q_{m,k+1})} + \frac{1}{2}q_{m,k+1}{r_2}^2$
$q_{m,k+1} > \frac{sg - R_{m,k}}{r_1 + g}$	$\frac{1}{2}(2R_{m,k} + q_{m,k+1}C)C - \frac{1}{2}(2C - g - 2r_1)sg$

Table 2. Analytical formulae for $d_{m,k+1}^{(2)}$.

Conditions	$\boldsymbol{d_{m,k+1}^{(2)}}$
$q_{m,k+1} \le \frac{sg_1 - R_{m,k}}{g_1}, q_{m,k+1} \le \frac{sg_2}{r + g_2}$	$\frac{{R_{m,k}}^2 + sq_{m,k+1}r^2}{2(s - q_{m,k+1})}$
$q_{m,k+1} \le \frac{sg_1 - R_{m,k}}{g_1}, q_{m,k+1} > \frac{sg_2}{r + g_2}$	$\frac{{R_{m,k}}^2}{2(s - q_{m,k+1})} + \frac{q_{m,k+1}r^2}{2} + \frac{g_2(2q_{m,k+1}r + q_{m,k+1}g_2 - sg_2)}{2}$
$q_{m,k+1} > \frac{sg_1 - R_{m,k}}{g_1}, q_{m,k+1} \le \frac{sg - R_{m,k}}{C}$	$\frac{(2R_{m,k} + q_{m,k+1}g_1 - sg_1)g_1}{2} + r\left(R_{m,k} + q_{m,k+1}g_1 + \frac{1}{2}q_{m,k+1}r - sg_1\right) + \frac{(R_{m,k} + q_{m,k+1}g_1 - sg_1 + q_{m,k+1}r)^2}{2(s - q_{m,k+1})}$
$q_{m,k+1} > \frac{sg_1 - R_{m,k}}{g_1}, q_{m,k+1} > \frac{sg - R_{m,k}}{C}$	$\frac{(2R_{m,k} + q_{m,k+1}g_1 - sg_1)g_1}{2}$

$$+r\left(R_{m,k}+q_{m,k+1}g_1+\frac{1}{2}q_{m,k+1}r-sg_1\right)$$
$$+\frac{1}{2}g_2\left[2R_{m,k}+q_{m,k+1}(r+g_1+C)\right.$$
$$\left.-s(2g_1+g_2)\right]$$

$$C=\frac{1}{\zeta}. \tag{4}$$

$$g=\frac{\phi_i+\zeta}{\zeta},\forall\, i\in[1,N_T]. \tag{5}$$

$$r_1=\frac{\theta_i}{\zeta},\forall\, i\in[1,N_T]. \tag{6}$$

$$r_2=\frac{1-\theta_i-\phi_i-\zeta}{\zeta},\forall\, i\in[1,N_T]. \tag{7}$$

$$r=\frac{1-\phi_i-\zeta}{\zeta},\forall\, i\in[1,N_T]. \tag{8}$$

$$g_1=\frac{\theta_i+\phi_i+\zeta-1}{\zeta},\forall\, i\in[1,N_T]. \tag{9}$$

$$g_2=\frac{1-\theta_i}{\zeta},\forall\, i\in[1,N_T]. \tag{10}$$

Proof. Under a uniform vehicle arrival pattern, traffic delay can be represented as the area of the polygon formed between the cumulative arrival and departure lines, as illustrated in Figure 3. Different demand levels form distinct polygons. This determines the various conditions and their corresponding formulae for estimating traffic delays, as shown in Tables 1 and 2. The detailed proof is provided in **Appendix C**.

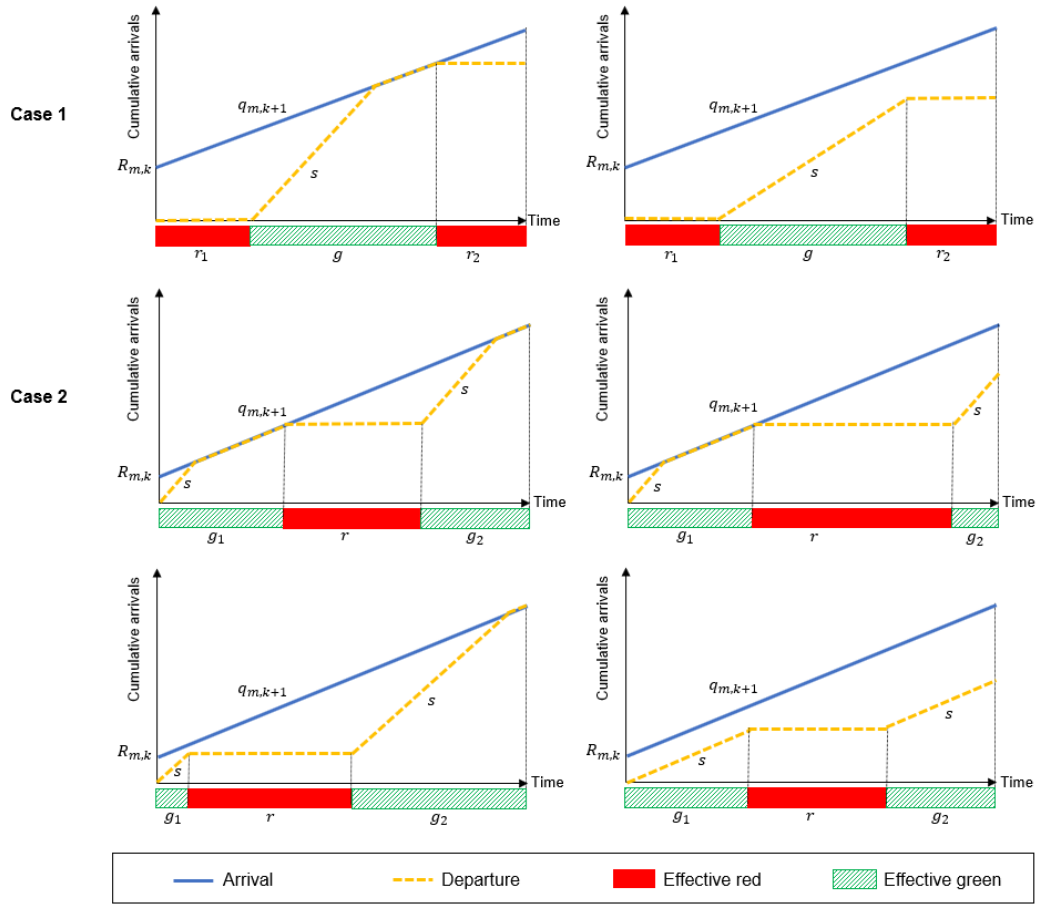


Figure 3. Estimation of traffic delays for lane m in cycle $k+1$. Cases 1 and 2 represent a signal group ending with an effective red and green, respectively.

■

Proposition 1 provides a generic and analytical model for estimating traffic delay under any signal plan. However, it only focuses on the traffic delay in cycle $k+1$ and fails to consider the potential impact caused by holding vehicles at the end of cycle $k+1$. In an ideal scenario, an effective signal plan should ensure that all arrivals are cleared within the same cycle, avoiding holding vehicles at the end of the green period. Failure to achieve this could result in significant delays for holding vehicles and subsequent arrivals, potentially causing congestion propagation in the local transportation system, particularly in situations with temporarily high V/C ratios. **Proposition 2** is introduced to model such consequential delays.

Proposition 2. Under the assumption that the temporarily high arrival rate returns to the average arrival rate $\bar{q}_{m,k}$ after cycle $k+1$, the consequential delay for any lane m in cycle $k+1$, $d^c_{m,k+1}$, can be estimated as

$$d^c_{m,k+1} = \gamma_1 {R_{m,k+1}}'^2 + \gamma_2 {R_{m,k+1}}', \forall m \in [1, l], \tag{11}$$

where

$$\gamma_1 = \frac{1}{l}\sum_{m=1}^{l}\frac{C_0}{2\left(sg_{m,0} - C_0\bar{q}_{m,k}\right)}; \tag{12}$$

$$\gamma_2 = \frac{1}{l}\sum_{m=1}^{l}\frac{r_{m,0}s}{2(s-\bar{q}_{m,k})}; \tag{13}$$

$R_{m,k+1}{}'$ represents the projected number of holding vehicles at the end of the nearest green period, and C_0, $g_{m,0}$, and $r_{m,0}$ respectively denote the cycle length, effective green duration, and effective red duration of a selected signal plan used for analysis. $R_{m,k+1}{}'$ can be estimated as either $R_{m,k+1}^{(1)}{}'$ or $R_{m,k+1}^{(2)}{}'$ depending on whether the lane is controlled by a signal group ending with effective red or green, as detailed in Tables 3 and 4.

Table 3. Analytical formulae for $R_{m,k+1}^{(1)}{}'$.

Conditions	$\boldsymbol{R_{m,k+1}^{(1)}{}'}$
$q_{m,k+1} \le \frac{sg - R_{m,k}}{r_1 + g}, \bar{q}_{m,k} > \frac{sg - q_{m,k+1}r_2}{r_1 + g}$	$q_{m,k+1}r_2 + \bar{q}_{m,k}(r_1 + g) - sg$
$q_{m,k+1} > \frac{sg - R_{m,k}}{r_1 + g}, \bar{q}_{m,k} > \frac{2sg - R_{m,k} - q_{m,k+1}C}{r_1 + g}$	$R_{m,k} + q_{m,k+1}C + \bar{q}_{m,k}(r_1 + g) - 2sg$
Otherwise	0

Table 4 Analytical formulae for $R_{m,k+1}^{(2)}{}'$.

Conditions	$\boldsymbol{R_{m,k+1}^{(2)}{}'}$
$q_{m,k+1} \le \frac{sg_1 - R_{m,k}}{g_1}, q_{m,k+1} > \frac{sg_2}{r + g_2}$, $\bar{q}_{m,k} > \frac{sg - q_{m,k+1}(r + g_2)}{g_1}$	$q_{m,k+1}(r + g_2) - sg + \bar{q}_{m,k}g_1$
$q_{m,k+1} > \frac{sg_1 - R_{m,k}}{g_1}, q_{m,k+1} > \frac{sg - R_{m,k}}{C}$, $\bar{q}_{m,k} > \frac{s(g_1 + g) - R_{m,k} - q_{m,k+1}C}{g_1}$	$R_{m,k} + q_{m,k+1}C - sg + (\bar{q}_{m,k} - s)g_1$
Otherwise	0

Proof. The existence of holding vehicles will delay all newly arrived vehicles owing to the first-come-first-served nature of the transportation system, resulting in extra delay except for the regular delay experienced by a system without holding vehicles. This additional delay is estimated as the consequential delay. A detailed proof is presented in **Appendix D**.

■

Proposition 2 reveals a quadratic relationship between $d^c_{m,k+1}$ and ${R_{m,k+1}}'$. To determine γ_1 and γ_2, a straightforward yet effective approach is to use the optimal fixed-time signal plan derived from average traffic demands as the selected analysis signal plan. With both **Propositions 1** and **2**, the total junction delay in cycle $k+1$ can be explicitly expressed as

$$D = f\left(\theta_1, \dots, \theta_i, \dots, \theta_{N_T+N_P}, \phi_1, \dots, \phi_i, \dots, \phi_{N_T+N_P}, \zeta \middle| p_{1,k}, \dots, p_{l,k}\right) = \sum_{m=1}^{l} \left(d_{m,k+1} + d^c_{m,k+1}\right). \quad (14)$$

3.2. Constraints for signal optimization

This subsection summarizes a set of constraints for signal optimization. First, the definitions of the control variables imply the following constraints:

$$0 \le \theta_i \le 1, \forall i \in [1, N_T + N_P]; \quad (15)$$

$$0 \le \phi_i \le 1, \forall i \in [1, N_T + N_P]. \quad (16)$$

Considering the maximum cycle length C_{max} gives

$$\zeta \ge \frac{1}{C_{max}}. \quad (17)$$

For traffic groups, the minimum signal group duration is denoted as g_T, meaning that

$$\phi_i \ge g_T \zeta, \forall i \in [1, N_T]. \quad (18)$$

For pedestrian groups,

$$\phi_i \ge g_P \zeta, \forall i \in [N_T + 1, N_T + N_P], \quad (19)$$

where the minimum signal group duration, g_P, is set to $\rho w / v_P$; $\rho = 1.5$ represents the provision being 50% higher than the minimum curb-to-curb crossing time; w is the width of the crossing in m; and $v_P = 1.2$ is the average walking speed in m/s. For safety, the clearance time constraints must be satisfied:

$$\theta_i + \phi_i + t_{ij}\zeta \le \theta_j + \Omega_{ij}, \forall (i,j) \in \mathbf{\Psi}. \quad (20)$$

In addition, in practical scenarios, the start or end of green for certain groups can be further constrained as follows:

$$\theta_i = \Delta_i \zeta, \forall i \in \mathbf{X_1}, \quad (21)$$

$$\theta_i + \phi_i \ge \sigma_i \zeta, \forall i \in \mathbf{X_2}, \quad (22)$$

where Δ_i and σ_i are user-defined parameters and $\mathbf{X_1}$ and $\mathbf{X_2}$ represent sets of groups that need to constrain the start and end of green, respectively. For instance, Eq. (21) can be used to set the start of green for a certain group as the beginning of a cycle. Equation (22) can be used to allocate computation time for optimization purposes.

3.3. Deterministic penetration rate control model

The objective function of the CVASC framework can be formulated as the minimization of the total junction delay, D, which is a function of the CV penetration rates. Given the uncertainty in the CV penetration rates, $p_{m,k}, \forall m \in [1, l]$, the objective function is also subject to uncertainty. One

straightforward approach for evaluating D involves disregarding the uncertainty in $p_{m,k}$ and replacing it with the estimated average and fixed CV penetration rate $\bar{p}_{m,k}$. Treating all of the input parameters as constants converts D into a deterministic model. Minimizing this deterministic delay, subject to the optimization constraints outlined in Section 3.2, results in the following formulation of the DPRC model:

$$\begin{aligned} &\min_{\{\theta_i,\phi_i,i\in[1,N_T+N_P]\},\zeta} J_D = D, \\ &s.t.\, D = f\left(\theta_1,\dots,\theta_i,\dots,\theta_{N_T+N_P},\phi_1,\dots,\phi_i,\dots,\phi_{N_T+N_P},\zeta\middle|\bar{p}_{1,k},\dots,\bar{p}_{m,k},\dots,\bar{p}_{l,k}\right), \\ &\quad \text{and Eqs. (15) to (22).} \end{aligned} \tag{23}$$

Given the analytical and deterministic form of J_D, substituting the analysis signal plan of interest into the function f provides a straightforward evaluation of the total junction delay. Furthermore, the gradient of J_D is given as follows:

$$\frac{\partial J_D}{\partial x} = \sum_{m=1}^{l}\left(\frac{\partial d_{m,k+1}}{\partial x} + \frac{\partial d^c_{m,k+1}}{\partial x}\right), \forall x \in \mathbb{V}, \tag{24}$$

where $\mathbb{V} = \left\{\theta_1,\dots,\theta_i,\dots,\theta_{N_T+N_P},\phi_1,\dots,\phi_i,\dots,\phi_{N_T+N_P},\zeta\right\}$ is the control variable set. Thus, the problem reduces to the derivation of the gradients of $d_{m,k+1}$ and $d^c_{m,k+1}$, which are respectively presented in **Appendices E** and **F**. Substituting the gradients of $d_{m,k+1}$ and $d^c_{m,k+1}$ into Eq. (24) provides the gradient of J_D.

3.4. Stochastic penetration rate control model

While the above deterministic model is straightforward, it does not consider the uncertainty in the CV penetration rate, and hence the uncertainty in vehicle arrival rate. In the context of the stochastic and dynamic nature of transportation systems, optimizing the system without accounting for this uncertainty may result in suboptimal solutions and unsatisfactory performance. To potentially achieve optimal solutions in adaptive signal control, incorporating the uncertainty of CV penetration rates in the estimation of D is imperative. As $p_{m,k}$ ranges from 0 to 1 by definition, it is assumed to follow a beta distribution with a mean of $\bar{p}_{m,k}$ and a variance of $Var(p_{m,k})$ (Jia et al., 2023, 2024a, 2024b). Consequently, the problem is reformulated into the following SPRC model:

$$\begin{aligned} &\min_{\{\theta_i,\phi_i,i\in[1,N_T+N_P]\},\zeta} J_S = E(D) + \omega\sqrt{Var(D)}, \\ &s.t.\, D = f\left(\theta_1,\dots,\theta_i,\dots,\theta_{N_T+N_P},\phi_1,\dots,\phi_i,\dots,\phi_{N_T+N_P},\zeta\middle|p_{1,k},\dots,p_{m,k},\dots,p_{l,k}\right) \\ &\quad \text{and Eqs. (15) to (22),} \end{aligned} \tag{25}$$

where ω is a user-defined parameter representing the trade-off between efficiency and stability.

$$p_{m,k} \sim Beta\left(\left[\frac{\bar{p}_{m,k}(1-\bar{p}_{m,k})}{Var(p_{m,k})} - 1\right]\bar{p}_{m,k}, \left[\frac{\bar{p}_{m,k}(1-\bar{p}_{m,k})}{Var(p_{m,k})} - 1\right](1-\bar{p}_{m,k})\right), \forall m \in [1,l]. \tag{26}$$

The inclusion of parameters with uncertainties in the objective function presents challenges in both evaluation and optimization. While the MCS method is a convenient option, it is time-consuming

and impractical for real-time applications. Moreover, obtaining the gradient of this objective function, which is essential for developing high-efficiency gradient-based solution algorithms, is an even more challenging task.

4. Analytical Stochastic Delay Model

This section introduces the analytical stochastic delay (ASD) model, which is designed to accurately estimate the stochastic quantity J_S and its gradient using efficient analytical models. Before the estimations of J_S and its gradient are presented, the independence of $p_{m,k}$, $\forall m \in [1, l]$, is established, as it is essential for the subsequent derivations.

4.1 Independence of $p_{m,k}$

4.1.1 Model derivation

The independence of $p_{m,k}$, $\forall m \in [1, l]$, is stated in **Lemma 1**.

Lemma 1. With the assumption that a super source demand q follows a Poisson distribution, i.e., $q \sim Pois(\lambda)$, which generates demands for all lanes, each vehicle is randomly determined to be a CV or a non-CV according to an underlying CV penetration rate. Additionally, each vehicle has an equal probability α_m of selecting lane m, where $\sum_{m=1}^{l} \alpha_m = 1$. Furthermore, the random variables $p_{m,k}$, $\forall m \in [1, l]$, are independent of each other.

Proof. Detailed proof of **Lemma 1** is provided in **Appendix G**.

■

4.1.2. Numerical experiments

To examine the independence of $p_{m,k}$, a series of numerical experiments is conducted. An approach with two arbitrary lanes connected to an intersection is considered. The total demand in each cycle is assumed to follow a Poisson distribution with a mean arrival of λ, i.e., $Pois(\lambda)$. Every vehicle is randomly assigned to be either a CV or a non-CV according to the underlying CV penetration rate p. In addition, each vehicle has an equal probability of selecting either lane 1 or lane 2, as represented by α_1 and α_2, respectively, where $\alpha_1 + \alpha_2 = 1$. For simplicity, a vertical queue assumption and a 50% red and 50% green signal plan are adopted. Various combinations of λ, α_1, and p are tested. For each combination, 10,000 cycles are simulated. In each cycle, the SSDPRE method is applied separately to the two lanes. This process generates 10,000 samples of $p_{m,k}$ for each lane. The correlation coefficient between $p_{1,k}$ and $p_{2,k}$ is then evaluated according to these samples. The correlation coefficients between $p_{1,k}$ and $p_{2,k}$ are very close to 0 across different combinations of λ, α_1, and p. Thus, the numerical experiments validate **Lemma 1**. Detailed analysis and results can be found in **Appendix H**.

4.2. Estimation of J_S

4.2.1. Model derivation

The objective function J_S in the SPRC model relies on both $E(D)$ and $Var(D)$, where D is dependent on uncertain parameters, $p_{m,k}$, $\forall m \in [1, l]$. Leveraging the independence property of $p_{m,k}$, **Proposition 3** is introduced for efficient estimations of $E(D)$ and $Var(D)$.

Proposition 3. For any given signal plan

$$\left(\theta_1 = \tilde{\theta}_1, \dots, \theta_i = \tilde{\theta}_i, \dots, \theta_{N_T+N_P} = \tilde{\theta}_{N_T+N_P}, \phi_1 = \tilde{\phi}_1, \dots, \phi_i = \tilde{\phi}_i, \dots, \phi_{N_T+N_P} = \tilde{\phi}_{N_T+N_P}, \zeta = \tilde{\zeta}\right), \quad (27)$$

with a stochastic model featuring distributed and independent parameters $p_{m,k}$, $\forall m \in [1, l]$,

$$\widetilde{D} = f\left(\tilde{\theta}_1, \dots, \tilde{\theta}_i, \dots, \tilde{\theta}_{N_T+N_P}, \tilde{\phi}_1, \dots, \tilde{\phi}_i, \dots, \tilde{\phi}_{N_T+N_P}, \tilde{\zeta} \middle| p_{1,k}, \dots, p_{m,k}, \dots, p_{l,k}\right), \quad (28)$$

and a K^{th}-order gPCE for $\widetilde{D}$,

$$\widetilde{D} = \sum_{|\boldsymbol{k}| \le K} c_{\boldsymbol{k}} \, \Phi_{\boldsymbol{k}}(\boldsymbol{Z}), \boldsymbol{Z} = (Z_1, \dots, Z_m, \dots, Z_l), \quad (29)$$

$E(\widetilde{D})$ and $Var(\widetilde{D})$ can be estimated as follows:

$$E(\widetilde{D}) = c_{\boldsymbol{0}}, \quad (30)$$

$$Var(\widetilde{D}) = \sum_{0 < |\boldsymbol{k}| \le \mathrm{K}} c_{\boldsymbol{k}}^2, \quad (31)$$

where $\{Z_1, \dots, Z_m, \dots, Z_l\}$ is a set of independent random variables determined by the original random variables $\{p_{1,k}, \dots, p_{m,k}, \dots, p_{l,k}\}$, $\{\Phi_{\boldsymbol{k}}(\boldsymbol{Z}) = \prod_{m=1}^{l} \Phi_{k_m}(Z_m), \forall |\boldsymbol{k}| \le K\}$ is a set of multivariate orthonormal polynomials, $\{\Phi_{k_m}(Z_m), \forall k_m \le K\}, \forall m \in [1, l]$ are sets of univariate orthonormal polynomials determined by the distribution of $Z_m, \forall m \in [1, l]$, $c_{\boldsymbol{k}}, \forall |\boldsymbol{k}| \le K$ is the coefficient of $\Phi_{\boldsymbol{k}}(\boldsymbol{Z})$, $\boldsymbol{k} = (k_1, k_2, \dots, k_l)$ is a multi-index with $|\boldsymbol{k}| = \sum_{j=1}^{l} k_j$, and K is the maximum order of gPCE.

Proof. Detailed proof of **Proposition 3** is provided in **Appendix I**.

■

Proposition 3 offers a fully analytical and efficient method to evaluate $E(D)$ and $Var(D)$ according to gPCE. Knowing the coefficients of gPCE is essential for evaluating $E(D)$ and $Var(D)$. **Proposition 4** is proposed for estimating these essential coefficients.

Proposition 4. Given the gPCE for $\widetilde{D}$, the set of ξ samples for $p_{m,k}, \forall m \in [1, l]$, $\left\{\boldsymbol{p}^{(i)} = \left(p_{1,k}^{(i)}, \dots, p_{l,k}^{(i)}\right), \forall i \in [1, \xi]\right\}$, the set of associated ξ samples for $\boldsymbol{Z}$, $\{\boldsymbol{Z}^{(i)}, \forall i \in [1, \xi]\}$, and the total number of cases for $|\boldsymbol{k}| \le K$ being $\mathbb{Q}$, the coefficients $c_{\boldsymbol{k}}, \forall |\boldsymbol{k}| \le K$ can be analytically estimated as follows:

$$\boldsymbol{c} = (\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{A}\boldsymbol{b}, \quad (32)$$

where

$$\boldsymbol{c} = \left[c_{\boldsymbol{k}_1}, \dots, c_{\boldsymbol{k}_{\mathbb{Q}}}\right]^{\mathrm{T}}, \mathbb{Q} = \binom{l+K}{K}, \tag{33}$$

$$\boldsymbol{A} = \begin{bmatrix} \Phi_{\boldsymbol{k}_1}\left(\boldsymbol{Z}^{(1)}\right) & \cdots & \Phi_{\boldsymbol{k}_1}\left(\boldsymbol{Z}^{(\xi)}\right) \\ \vdots & \ddots & \vdots \\ \Phi_{\boldsymbol{k}_{\mathbb{Q}}}\left(\boldsymbol{Z}^{(1)}\right) & \cdots & \Phi_{\boldsymbol{k}_{\mathbb{Q}}}\left(\boldsymbol{Z}^{(\xi)}\right) \end{bmatrix}, \tag{34}$$

$$\boldsymbol{b} = \begin{bmatrix} f\left(\tilde{\theta}_1, \dots, \tilde{\theta}_i, \dots, \tilde{\theta}_{N_T+N_P}, \tilde{\phi}_1, \dots, \tilde{\phi}_i, \dots, \tilde{\phi}_{N_T+N_P}, \tilde{\zeta} | \boldsymbol{p}^{(1)}\right) \\ \dots \\ f\left(\tilde{\theta}_1, \dots, \tilde{\theta}_i, \dots, \tilde{\theta}_{N_T+N_P}, \tilde{\phi}_1, \dots, \tilde{\phi}_i, \dots, \tilde{\phi}_{N_T+N_P}, \tilde{\zeta} | \boldsymbol{p}^{(\xi)}\right) \end{bmatrix}. \tag{35}$$

Proof. Detailed proof of **Proposition 4** is provided in **Appendix J**.

■

With **Proposition 4**, the coefficients $c_{\boldsymbol{k}}, \forall |\boldsymbol{k}| \leq K$ can be efficiently obtained using only a small number of samples. Typically, the number of samples ξ is set to $\mu\mathbb{Q}$, where μ represents the oversampling rate, and it is commonly set to 2 or 3 ($\mu = 2$ is used in this paper). To ensure the reproducibility of the experiment and improve convergence speed in coefficient estimation, the samples $\boldsymbol{p}^{(i)} = \left(p_{1,k}^{(i)}, p_{2,k}^{(i)}, \dots, p_{l,k}^{(i)}\right), \forall i \in [1, \xi]$ are generated using quasi-random and low-discrepancy Sobol sequences. Notably, the selected distribution for $Z_m, \forall m \in [1, l]$ and the associated orthonormal polynomials are dependent on the original distribution $p_{m,k}, \forall m \in [1, l]$. In this study, $p_{m,k}$ is assumed to follow a beta distribution; hence, Jacobi polynomials are used. Given that Jacobi polynomials are formulated using Rodrigues's formula, the support of the associated weighting function is $[-1,1]$ rather than the standard beta distribution defined on $[0,1]$. Therefore, necessary transformations are required between $p_{m,k}$ and Z_m. Detailed derivation is provided in **Appendix K**.

4.2.2. Numerical experiments

This subsection presents a series of numerical experiments aimed at comparing the efficiency of the MCS method and **Propositions 3** and **4** in evaluating $E(D)$ and $Var(D)$. Consider the real-world four-arm intersection depicted in Figure 5 in Section 6. Given a total of 13 approaching lanes connected to the intersection, the total junction delay D depends not only on parameters such as the signal plan parameters but also on the 13 uncertain CV penetration rate parameters $p_{m,k}$, $\forall m \in [1,13]$. These random variables greatly complicate the task of evaluating $E(D)$ and $Var(D)$. The MCS method estimates $E(D)$ and $Var(D)$ by sampling uncertain parameters from the assumed distributions. Despite its simplicity, this method typically requires millions of samples to converge and may lack reproducibility. In contrast, the proposed **Proposition 4** requires only a small number of samples to estimate the essential inputs of **Proposition 3**, making it much more efficient.

The numerical experiment results indicate that when the MCS method is used, the estimates converge as the number of samples gradually increases from 10 to 10^7 at a growth rate of 10. With a negligible difference observed between the estimates obtained using sample sizes of 10^6 and 10^7, the estimates derived from the MCS method with 10^7 samples are considered the ground-truth values,

requiring a computation time of ~24 s. In contrast, when K in **Proposition 4** is set to 1, only 28 samples are required, and **Proposition 3** achieves absolute percentage errors (APEs) of 0% and 28.9% in estimating $E(D)$ and $Var(D)$, respectively, with negligible computational cost. Increasing the approximation order K to 2 and 3 reduces the APE of estimated $Var(D)$ to 3.2% and 0.09%, respectively, while the APE of estimated $E(D)$ is maintained within 0.07%. Considering both approximation accuracy and computational complexity, $K = 2$ is chosen. With this choice, only as few as 210 samples are required for estimation, and thus the approach is several orders of magnitude more efficient than the MCS method. The computation time for **Propositions 3** and **4** is only 0.08 s, approximately 300 times more efficient than the MCS method. These results consistently indicate the superiority and efficiency of **Propositions 3** and **4**. More details of the numerical experiment are presented in **Appendix L**.

4.3. Gradient estimation of J_S

4.3.1. Model derivation

According to Eq. (32), the gradient of J_S is expressed as

$$\frac{\partial J_S}{\partial x} = \frac{\partial E(D)}{\partial x} + \frac{\omega}{2\sqrt{Var(D)}} \frac{\partial Var(D)}{\partial x}, \forall x \in \mathbb{V}. \tag{36}$$

Therefore, estimating the gradients of $E(D)$ and $Var(D)$ is imperative; the estimation is detailed in **Proposition 5**.

Proposition 5. Given D as defined in Eq. (25), $c_{\boldsymbol{k}_1} = c_{\mathbf{0}}$ in Eq. (33),

$$\mathbb{A} = (\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{A} = \begin{bmatrix} a_{11} & \cdots & a_{1\xi} \\ \vdots & \ddots & \vdots \\ a_{\mathbb{Q}1} & \cdots & a_{\mathbb{Q}\xi} \end{bmatrix} = \begin{bmatrix} \mathbb{A}_1 \\ \cdots \\ \mathbb{A}_{\mathbb{Q}} \end{bmatrix}, \tag{37}$$

and the other notation consistent with that in **Proposition 4**, the gradients of $E(D)$ and $Var(D)$ are respectively expressed as follows:

$$\frac{\partial E(D)}{\partial x} = \mathbb{A}_1 \boldsymbol{b}', \forall x \in \mathbb{V}, \tag{38}$$

$$\frac{\partial Var(D)}{\partial x} = \sum_{i=2}^{\mathbb{Q}} 2c_{\boldsymbol{k}_i} \mathbb{A}_i \boldsymbol{b}', \forall x \in \mathbb{V}, \tag{39}$$

where

$$\boldsymbol{b}' = \begin{bmatrix} f_x'\left(\theta_1, \dots, \theta_i, \dots, \theta_{N_T+N_P}, \phi_1, \dots, \phi_i, \dots, \phi_{N_T+N_P}, \zeta | \boldsymbol{p}^{(1)}\right) \\ \cdots \\ f_x'\left(\theta_1, \dots, \theta_i, \dots, \theta_{N_T+N_P}, \phi_1, \dots, \phi_i, \dots, \phi_{N_T+N_P}, \zeta | \boldsymbol{p}^{(\xi)}\right) \end{bmatrix}. \tag{40}$$

Proof. Detailed proof of **Proposition 5** is provided in **Appendix M**.

■

Proposition 5 completes the ASD model, indicating that the gradients of $E(D)$ and $Var(D)$ can be theoretically estimated according to the gPCE and gradients of the corresponding deterministic models. This crucial finding enables the development of efficient gradient-based algorithms for addressing stochastic optimization problems.

4.3.2. Numerical experiments

This subsection validates the effectiveness of **Proposition 5** in estimating the gradients of $E(D)$ and $Var(D)$ through a series of numerical experiments. The experiments are conducted using the settings and parameters outlined in Section 4.2.2. The gradients of $E(D)$ and $Var(D)$ are first estimated through a numerical differentiation method, and the values serve as the ground truth for evaluation. Subsequently, with $K = 2$, the analytical models proposed in **Proposition 5** are used to efficiently estimate the gradients using only 210 samples.

The results of estimates obtained using **Proposition 5** indicate that the maximum APE, average APE, and variance in APE of the estimated partial derivatives of $E(D)$ are 0.99%, 0.14%, and 0.08%, respectively. These near-zero metrics confirm the accuracy of **Proposition 5** in estimating the gradient of $E(D)$. Similarly, the maximum APE, average APE, and variance in APE of the estimated partial derivatives of $Var(D)$ are 4.21%, 2.56%, and 1.51%, respectively. Given the inherently greater complexity of estimating $Var(D)$ and its gradient, these errors are slightly larger than those of $E(D)$, as expected. Nonetheless, the maximum APE among all partial derivatives remains below 5%, with an average APE of only ~3%. Moreover, the computation time for **Proposition 5** is only 0.08 s, significantly shorter than that required by the numerical differentiation method (525.42 s). These results consistently demonstrate the superiority of **Proposition 5**. Further details are provided in **Appendix N**.

5. Gradient-guided Golden Section Search Algorithm

The gradient-based method optimizes the signals along the most efficient direction but may become trapped in a local minimum. The golden section search method searches for solutions along a specific line spanning the entire feasible region, but randomly selecting the search line can significantly degrade efficiency. This section proposes an efficient gradient-guided golden section search (G^3S^2) algorithm that leverages the strengths of both approaches to dynamically optimize the signal plan.

Let $J(\cdot)$ represent the objective function, which can be either J_D or J_S. The G^3S^2 algorithm is tailored for sequential optimization problems and uses the optimized signal plan for cycle k (denoted as $\mathbb{V}^{(k,*)}$) as the initial solution to optimize the signal plan for cycle $k+1$, as depicted in **Algorithm 1**. For the initial cycle when $k = 0$, $\mathbb{V}^{(0,*)}$ can be either randomly generated or manually set according to prior knowledge. In step 1, the random signal generation algorithm, denoted by $\mathbb{G}\left(\mathbb{V}^{(k,*)}, N_R\right)$ and detailed in **Algorithm 2**, is applied. The algorithm uses the optimized signal plan for cycle k, $\mathbb{V}^{(k,*)}$ as input to generate N_R additional random signals. This process helps to avoid poor local minima and aims to search for better solutions. Subsequently, the G^3S^2 procedure is applied separately to $N_R + 1$ initial

signal plans, resulting in $N_R + 1$ candidate signal plans. The optimized signal plan for cycle $k+1$ is selected as the best signal plan among the candidate signal plans according to their objective function values. Rather than randomly selecting a direction for the golden section search, the gradient direction is used to accelerate convergence, as demonstrated in steps 5 and 6 in **Algorithm 1**. Additionally, linear programming (LP) problems are involved in step 6 of **Algorithm 1** and step 4 of **Algorithm 2**, which can be efficiently solved through the simplex method. Applying the proposed algorithms at the end of each cycle enables cycle-by-cycle adaptive signal control. Figure 4 illustrates the G^{3}S^2 algorithm.

Algorithm 1. Gradient-guided golden section search (G^{3}S^2) algorithm.

Input: Optimized signal plan for cycle k, $\mathbb{V}^{(k,*)} = \left(\theta_1^{(k,*)}, \dots, \theta_i^{(k,*)}, \dots \theta_{N_T+N_P}^{(k,*)}, \phi_1^{(k,*)}, \dots, \phi_i^{(k,*)}, \dots \phi_{N_T+N_P}^{(k,*)}, \zeta^{(k,*)}\right)$, where $k \geq 0$, N_R is the number of random signal plans, N_{GRA} is the maximum iterations of gradient computation, N_{GSS} is the maximum iterations of golden section search under each gradient direction, and $\mathbb{p} = \frac{\sqrt{5}-1}{2}$ is the golden section number.

Output: Optimized signal plan for cycle $k+1$, $\mathbb{V}^{(k+1,*)} = \left(\theta_0^{(k+1,*)}, \dots, \theta_i^{(k+1,*)}, \dots \theta_{N_T+N_P}^{(k+1,*)}, \phi_0^{(k+1,*)}, \dots, \phi_i^{(k+1,*)}, \dots, \phi_{N_T+N_P}^{(k+1,*)}, \zeta^{(k+1,*)}\right)$.

Initialization: $S_{D^*} \leftarrow \{\}$, $S_{\mathbb{V}^*} \leftarrow \{\}$

1: **for** $\widetilde{\mathbb{V}}$ in $\mathbb{G}\left(\mathbb{V}^{(k,*)}, N_R\right)$ **do**

2: $\quad \mathbb{V}^* \leftarrow \widetilde{\mathbb{V}}$

3: $\quad$ **for** $i = 1, 2, \dots, N_{GRA}$ **do**

4: $\quad\quad \widetilde{\mathbb{V}} \leftarrow \mathbb{V}^*$

5: $\quad\quad$ Compute gradient $\Delta\widetilde{\mathbb{V}} = \left(J'_{\theta_0}\big|_{\widetilde{\mathbb{V}}}, \dots, J'_{\theta_{N_T+N_P}}\Big|_{\widetilde{\mathbb{V}}}, J'_{\phi_0}\big|_{\widetilde{\mathbb{V}}}, \dots, J'_{\phi_{N_T+N_P}}\Big|_{\widetilde{\mathbb{V}}}, J'_{\zeta}\big|_{\widetilde{\mathbb{V}}}\right)$

6: $\quad\quad$ Find η_{max} by substituting $\mathbb{V} \leftarrow \widetilde{\mathbb{V}} - \eta\Delta\widetilde{\mathbb{V}}$ into Eqs. (15) to (22), i.e.,

$$\eta_{max} = \arg\max \eta$$

$$s.t. \tilde{\zeta} - \eta J'_{\zeta}\big|_{\widetilde{\mathbb{V}}} \geq \frac{1}{C_{max}}$$

$$0 \leq \tilde{\theta}_j - \eta\, J'_{\theta_j}\Big|_{\widetilde{\mathbb{V}}} \leq 1, \forall j \in [1, N_T + N_P]$$

$$0 \leq \tilde{\phi}_j - \eta\, J'_{\phi_j}\Big|_{\widetilde{\mathbb{V}}} \leq 1, \forall j \in [1, N_T + N_P]$$

$$\tilde{\phi}_j - \eta\, J'_{\phi_j}\Big|_{\widetilde{\mathbb{V}}} \geq g_T\left(\tilde{\zeta} - \eta J'_{\zeta}\big|_{\widetilde{\mathbb{V}}}\right), \forall j \in [1, N_T]$$

$$\tilde{\phi}_j - \eta\, J'_{\phi_j}\Big|_{\widetilde{\mathbb{V}}} \geq g_P\left(\tilde{\zeta} - \eta J'_{\zeta}\big|_{\widetilde{\mathbb{V}}}\right), \forall j \in [N_T + 1, N_T + N_P]$$

$$\tilde{\theta}_u - \eta J'_{\theta_u}\big|_{\widetilde{\mathbb{V}}} + \tilde{\phi}_u - \eta J'_{\phi_u}\big|_{\widetilde{\mathbb{V}}} + t_{uv}\left(\tilde{\zeta} - \eta J'_{\zeta}\big|_{\widetilde{\mathbb{V}}}\right) \leq \tilde{\theta}_v - \eta J'_{\theta_v}\big|_{\widetilde{\mathbb{V}}} + \Omega_{uv}, \forall (u, v) \in \Psi$$

$$\tilde{\theta}_j - \eta\, J'_{\theta_j}\Big|_{\widetilde{\mathbb{V}}} = \Delta_j\left(\tilde{\zeta} - \eta J'_{\zeta}\big|_{\widetilde{\mathbb{V}}}\right), \forall j \in \mathrm{X}_1$$

$\tilde{\theta}_j - \eta J'_{\theta_j}\big|_{\widetilde{\mathbb{V}}} + \tilde{\phi}_j - \eta J'_{\phi_j}\big|_{\widetilde{\mathbb{V}}} \geq \sigma_j\left(\tilde{\zeta} - \eta J'_{\zeta}\big|_{\widetilde{\mathbb{V}}}\right), \forall j \in \mathrm{X}_2$

$\eta_{min} = 0$

for $m = 1, 2, \ldots, N_{GSS}$ **do**

$\eta_1 \leftarrow \eta_{min} + (1 - \mathbb{P})(\eta_{max} - \eta_{min})$

$\eta_2 \leftarrow \eta_{min} + \mathbb{P}(\eta_{max} - \eta_{min})$

$\eta^* \leftarrow \arg\min J\left(\widetilde{\mathbb{V}} - \eta\nabla\widetilde{\mathbb{V}}\right)\ s.t. \eta \in [\eta_{min}, \eta_1, \eta_2, \eta_{max}]$

if η^* in $[\eta_{min}, \eta_1]$ **then**

$\eta_{max} \leftarrow \eta_2$

else

$\eta_{min} \leftarrow \eta_1$

$\mathbb{V}^* \leftarrow \widetilde{\mathbb{V}} - \eta^*\Delta\widetilde{\mathbb{V}}$

$D^* = J\left(\widetilde{\mathbb{V}} - \eta^*\Delta\widetilde{\mathbb{V}}\right)$

$S_{D^*} \leftarrow S_{D^*} + \{D^*\}$

$S_{\mathbb{V}^*} \leftarrow S_{\mathbb{V}^*} + \{\mathbb{V}^*\}$

$\mathbb{V}^{(k+1,*)} = S_{\mathbb{V}^*}[\arg\min(S_{D^*})]$

Algorithm 2. Random signal generation algorithm.

Input: Optimized signal plan for cycle k, $\mathbb{V}^{(k,*)} = \left(\theta_1^{(k,*)}, \ldots, \theta_i^{(k,*)}, \ldots \theta_{N_T+N_P}^{(k,*)}, \phi_1^{(k,*)}, \ldots, \phi_i^{(k,*)}, \ldots \phi_{N_T+N_P}^{(k,*)}, \zeta^{(k,*)}\right)$, where $k \geq 0$ and N_R is the number of random signal plans.

Output: A set of starting signal plans $S_{\mathbb{V}} = \mathbb{G}\left(\mathbb{V}^{(k,*)}, N_R\right)$.

Initialization: $S_{\mathbb{V}} \leftarrow \{\}$

$S_{\mathbb{V}} \leftarrow S_{\mathbb{V}} + \left\{\mathbb{V}^{(k,*)}\right\}$

for $i = 1,2, \ldots, N_R$ **do**

Randomly generate a direction $\left(\epsilon_1, \ldots, \epsilon_{2(N_T+N_P)+1}\right)$, $\epsilon_j \sim U(-1,1), \forall j \in [1,2(N_T + N_P) + 1]$.

Solve the following LP problems:

$$\eta_{min} = \arg\min \eta \ \ and \ \eta_{max} = \arg\max \eta$$

$$s.t. \zeta^{(k,*)} - \eta\epsilon_{2(N_T+N_P)+1} \geq \frac{1}{C_{max}}$$

$$0 \leq \theta_j^{(k,*)} - \eta\epsilon_j \leq 1, \forall j \in [1, N_T + N_P]$$

$$0 \leq \phi_j^{(k,*)} - \eta\epsilon_j \leq 1, \forall j \in [1, N_T + N_P]$$

$$\phi_j^{(k,*)} - \eta\epsilon_j \geq g_T\left(\zeta^{(k,*)} - \eta\epsilon_{2(N_T+N_P)+1}\right), \forall j \in [1, N_T]$$

$$\phi_j^{(k,*)} - \eta\epsilon_j \geq g_P\left(\zeta^{(k,*)} - \eta\epsilon_{2(N_T+N_P)+1}\right), \forall j \in [N_T + 1, N_T + N_P]$$

$$\theta_u^{(k,*)} - \eta\epsilon_u + \phi_u^{(k,*)} - \eta\epsilon_u + t_{uv}\left(\zeta^{(k,*)} - \eta\epsilon_{2(N_T+N_P)+1}\right) \le \theta_v^{(k,*)} - \eta\epsilon_v + \Omega_{uv}, \forall (u,v) \in \Psi$$

$$\theta_j^{(k,*)} - \eta\epsilon_j = \Delta_j\left(\zeta^{(k,*)} - \eta\epsilon_{2(N_T+N_P)+1}\right), \forall j \in \mathrm{X}_1$$

$$\theta_j^{(k,*)} - \eta\epsilon_j + \phi_j^{(k,*)} - \eta\epsilon_j \ge \sigma_j\left(\zeta^{(k,*)} - \eta\epsilon_{2(N_T+N_P)+1}\right), \forall j \in \mathrm{X}_2$$

5: Randomly generate a number $\vartheta \sim U(0,1)$

6: $\eta_{mid} = \eta_{min} + \vartheta(\eta_{max} - \eta_{min})$

7: $S_{\mathbb{V}} \leftarrow S_{\mathbb{V}} + \left\{\mathbb{V}^{(k,*)} - \eta_{mid}\left(\epsilon_1, \dots, \epsilon_{2(N_T+N_P)+1}\right)\right\}$

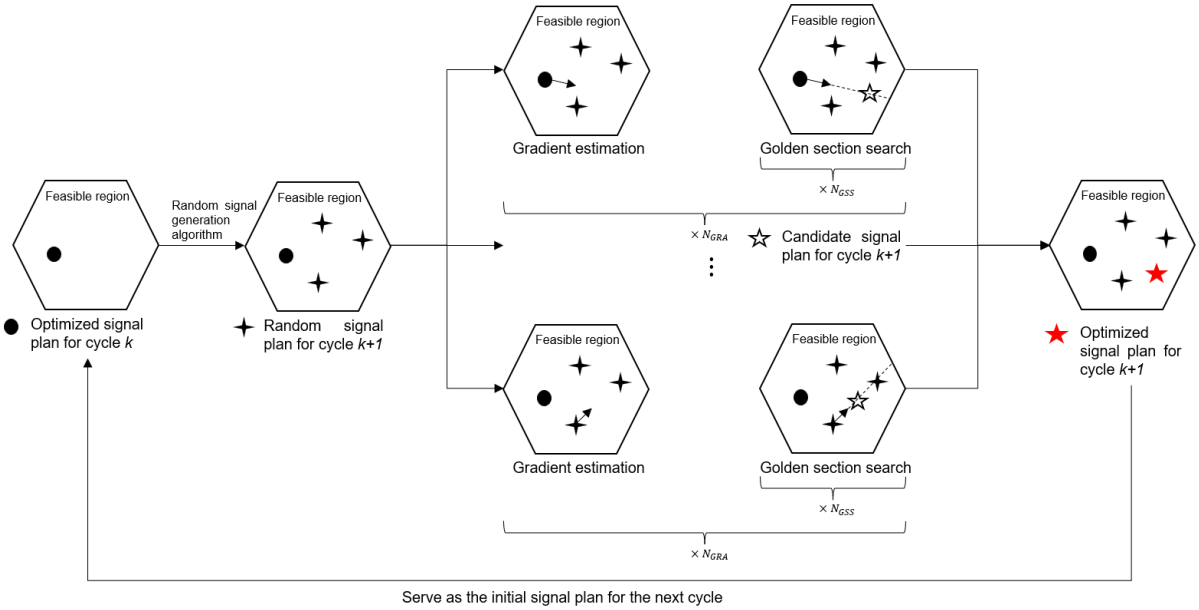


Figure 4. Illustration of the G^{3}S^2 algorithm.

6. Realistic Simulations

This section presents a comprehensive and realistic simulation study to demonstrate the applicability and efficiency of the proposed CVASC framework and its solution methods.

6.1. General settings

Simulations were performed using a VISSIM platform in a Windows 10 environment on a machine equipped with an Intel Core i7-10700 CPU. Intersection 1 in the next-generation simulation (NGSIM) dataset, a real-world intersection at the junction of Peachtree Street and 10th Street in Atlanta, Georgia, USA, was selected for the simulation study. The original layout of the intersection did not include any pedestrian crossings. However, given the crucial role that pedestrians play at intersections, many studies have specifically considered pedestrians in intersection control (e.g., Liang et al., 2020b; Yin et al., 2021). To demonstrate the versatility of the proposed framework, two pedestrian crossings were introduced. The modified intersection layout and the corresponding groups are illustrated in

Figures 5 and 6, respectively. Additionally, the actual demands and turning proportions between 12:45 and 13:00 on November 8, 2006 were extracted and are provided in Table O1 in **Appendix O**.

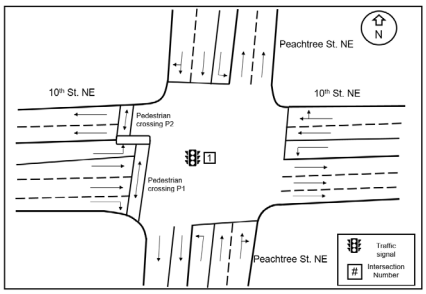


Figure 5. Modified Intersection 1.

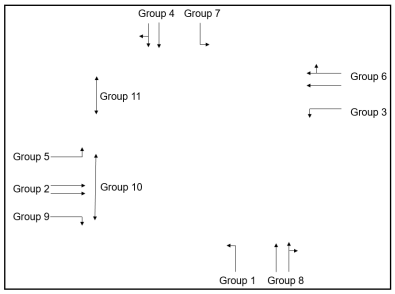


Figure 6. Corresponding signal groups.

Various simulation cases were constructed using different V/C ratios (i.e., 0.3, 0.5, and 0.7) and underlying CV penetration rates (i.e., 0.1, 0.4, and 0.7). For each simulation case, the study period was set to 2 h after a 30-min warm-up period. Vehicle arrivals followed Poisson distributions with constant average demands during the warm-up period, and the optimal fixed-time signal plan derived from the optimised signal capacity and delay (OSCADY) model was implemented. To simulate a realistic rush-hour traffic demand pattern, each approach was set to have a commonly used triangular demand profile, in which the demand starts and ends at 0.8 times the average demand, peaking at 1.2 times the average demand. That is, while the average demand for each approach varied over time, the mean of these average demands remained identical to the constant average demand used in the warm-up period. In addition, the triangular demand profile period aligned with the duration of the study period. The demand generated under each average demand followed a Poisson distribution. To test different V/C ratios, the average demands corresponding to the V/C ratios of interest (i.e., 0.3, 0.5, and 0.7) were derived by scaling the actual demands up or down. Furthermore, each vehicle had probabilities p and $1-p$ of being a CV and a non-CV, respectively, with different underlying CV penetration rates considered (i.e., 0.1, 0.4, and 0.7).

6.2. Signal control schemes

To critically evaluate the performance of the proposed control schemes, namely the DPRC and SPRC schemes, within the CVASC framework, the built-in vehicle-actuated control (VAC) scheme in VISSIM was chosen as the benchmark. VAC relies on downstream detectors placed on each lane to detect vehicle presence and adaptively adjusts signal timings based on real-time traffic data. Thus, as the VAC scheme uses full traffic information, it has demonstrated reliable and satisfactory performance and has been widely employed as a benchmark (Feng et al., 2015; Feng et al., 2018; Wang et al., 2021; Tan and Yang, 2024). Given that the VAC scheme operates without a fixed cycle length, the total junction delay (defined as the sum of all vehicle delays experienced during the study period) was selected as the evaluation metric for all signal control schemes under consideration.

In the VAC scheme, the minimum durations for traffic signal groups were set to 5 s, while the minimum durations for pedestrian signal groups 10 and 11 (depicted in Figures 5 and 6) were set to 18 s and 9 s, respectively, based on a lane width of 3.5 m. The maximum green and vehicle extension times for each signal group were determined by a simple grid search method with step sizes of 10 s and 2 s, respectively. For each combination of maximum green and vehicle extension times, a 2-h simulation was conducted to determine the resulting total junction delay, with the combination yielding the lowest total junction delay selected as the optimal parameters. Amber and red clearance times for all signal groups were set to 3 s and 2 s, respectively. The controller frequency was set to 2 Hz, enabling communication with VISSIM twice per second during the simulation. Following a 30-min warm-up period, the VAC scheme was implemented with optimal parameters in VISSIM to control all traffic signals over the 2-h study period. Vehicle delays recorded by VISSIM during this period were subsequently stored for evaluation.

In contrast, both DPRC and SPRC involved cycle-by-cycle adaptive signal control solely based on CV data. These schemes optimized the signal plan at the end of each cycle and implemented the optimized plan in the next cycle. The successor matrix and the clearance time matrix (Tables O2 and O3 in **Appendix O**) determined the group sequence and the clearance times, respectively. These matrices were essential parameters in DPRC and SPRC. Both schemes followed the same set of constraints outlined in Eqs. (15) to (22). Without loss of generality, the start of green of signal group 1 was set to the beginning of a cycle. As with VAC, the minimum durations for traffic signal groups were set to 5 s, while those for pedestrian signal groups 10 and 11 were set to 18 s and 9 s, respectively. C_{max} was set to 120 s. Considering the computational constraints in real-world applications, a 3 s buffer was reserved at the start of the signal plan, necessitating $\theta_9 + \phi_9 - 1 \geq 3\zeta$ given the defined intersection and signal groups. After the observation of several cycles, the saturation flow rate and the average effective vehicle length were determined as 2,268 veh/h and 6.44 m, respectively. The primary distinction between DPRC and SPRC lies in the treatment of the CV penetration rate uncertainty. DPRC does not consider this uncertainty, resulting in fixed vehicle arrival rates and deterministic control, while SPRC incorporates its uncertainty, and hence the uncertainty in vehicle arrival rates, leading to stochastic control. DPRC, SPRC with $\omega = 0$, and SPRC with $\omega = 20$ were considered in the simulation study. To implement the proposed G^3S^2 algorithm, the number of random signal plans explored in each optimization cycle, N_R, must be determined. A small N_R may result in the failure to capture potentially better solutions, while excessively large N_R values may yield only marginal improvement at a high computational cost. Preliminary experiments determined appropriate values for N_R as follows: 19 for DPRC, 29 for SPRC with $\omega = 0$, and 39 for SPRC with $\omega = 20$. Further details are provided in **Appendix P**. Additionally, both N_{GRA} and N_{GSS} in **Algorithm 1** were set to 5. The initial signal plan $\mathbb{V}^{(0,*)}$ for both DPRC and SPRC was configured as the optimal fixed-time signal plan to

ensure fair comparisons. Each simulation case consisted of a 30-min warm-up followed by a 2-h study period, during which individual vehicle delays were recorded for evaluation.

6.3. Results

Table 5 presents the performances of different signal control schemes at a low V/C ratio (0.3), with the last column indicating the percentage improvement in total junction delay of the proposed control schemes compared with the VAC scheme benchmark. Given the light traffic demand, the performances of the signal control schemes were not expected to vary significantly. The DPRC scheme, which does not account for CV penetration rate uncertainty, exhibited slightly inferior but comparable performance to the VAC scheme across different levels of CV penetration. Performance worsened as CV penetration rates decreased and improved as CV penetration rates increased. In contrast to the DPRC scheme, the SPRC scheme accounts for CV penetration rate uncertainty. The SPRC scheme with ω = 20, which optimized for both efficiency and stability, only partially considered traffic efficiency but demonstrated performance that was comparable to or slightly better than the performances of the DPRC and VAC schemes across different CV penetration rates. In contrast, the SPRC scheme with ω = 0, which optimized for only traffic efficiency, outperformed the VAC and DPRC schemes at all CV penetration rates (i.e., 0.1, 0.4, and 0.7). Similarly, performance improved with increased CV penetration rates. These results highlight the importance of considering CV penetration rate uncertainty in CV-based traffic signal control.

Table 5. Performance comparison of different signal control schemes at a V/C ratio of 0.3 with varying $\mathscr{p}$.

Scheme	$\mathscr{p}$	Incorporation of $Var(p_{m,k})$?	Total junction delay (s)	Improvement (%)
VAC	-	-	79,004	-
DPRC	0.1	×	79,376	-0.47
DPRC	0.4	×	79,016	-0.02
DPRC	0.7	×	79,007	-0.00
SPRC ($\omega = 20$)	0.1	✓	79,177	-0.22
SPRC ($\omega = 20$)	0.4	✓	79,018	-0.02
SPRC ($\omega = 20$)	0.7	✓	78,706	0.38
SPRC ($\omega = 0$)	0.1	✓	78,563	0.56
SPRC ($\omega = 0$)	0.4	✓	77,666	1.69
SPRC ($\omega = 0$)	0.7	✓	77,376	2.06

Table 6 presents the performances of different signal control schemes at a medium V/C ratio (0.5), and it can be seen that the patterns of performance are similar to those depicted in Table 5. The

performance of the DPRC scheme improved as CV penetration rates increased but remained slightly inferior to that of the VAC scheme across all CV penetration rates (0.1, 0.4, and 0.7). The SPRC scheme with ω = 20 only partially considered traffic efficiency and thus its total junction delay improvements were -2.97%, -0.47%, and 1.03% (compared with the VAC scheme) at CV penetration rates of 0.1, 0.4, and 0.7, respectively. In contrast, the SPRC scheme with ω = 0 fully optimized traffic efficiency and thus its total junction delay improvements were more significant, namely 2.99%, 5.36%, and 7.03% (compared with the VAC scheme) at CV penetration rates of 0.1, 0.4, and 0.7, respectively. Table 7 presents the performance of different signal control schemes at a high V/C ratio (0.7), and it can be seen that the DPRC scheme exhibited worse performance than the VAC scheme at CV penetration rates of 0.1 and 0.4 but better performance than the VAC scheme at a CV penetration rate of 0.7. The SPRC scheme with ω = 20 only partially considered traffic efficiency and thus did not demonstrate any improvements (compared with the VAC scheme) in total junction delay at CV penetration rates of 0.1, 0.4, and 0.7. However, the SPRC scheme with ω = 0 fully optimized traffic efficiency and thus outperformed the VAC scheme by 4.44% and 12.62% in total junction delay at CV penetration rates of 0.4 and 0.7, respectively. These findings suggest that across a wide range of CV penetration rates and V/C ratios, the proposed CV-based signal control schemes, which do not rely on detector data, generally perform similarly to or better than the VAC scheme, which requires complete traffic information from detectors. Furthermore, these findings underscore the significance and superiority of incorporating CV penetration rate uncertainty into CV-based traffic signal control.

Table 6. Performance comparison of different signal control schemes at a V/C ratio of 0.5 with varying $\mathscr{p}$.

Scheme	$\mathscr{p}$	Incorporation of $Var(p_{m,k})$?	Total junction delay (s)	Improvement (%)
VAC	-	-	146,607	-
DPRC	0.1	×	149,297	-1.83
DPRC	0.4	×	148,278	-1.14
DPRC	0.7	×	148,192	-1.08
SPRC ($\omega = 20$)	0.1	✓	150,960	-2.97
SPRC ($\omega = 20$)	0.4	✓	147,301	-0.47
SPRC ($\omega = 20$)	0.7	✓	145,093	1.03
SPRC ($\omega = 0$)	0.1	✓	142,226	2.99
SPRC ($\omega = 0$)	0.4	✓	138,745	5.36
SPRC ($\omega = 0$)	0.7	✓	136,301	7.03

Table 7. Performance comparison of different signal control schemes at a V/C ratio of 0.7 with varying $\mathscr{p}$.

Scheme	$\mathscr{p}$	Incorporation of $Var(p_{m,k})$?	Total junction delay (s)	Improvement (%)
VAC	-	-	329,837	-
DPRC	0.1	×	431,355	-30.78
DPRC	0.4	×	355,579	-7.80
DPRC	0.7	×	323,523	1.91
SPRC ($\omega = 20$)	0.1	✓	397,908	-20.64
SPRC ($\omega = 20$)	0.4	✓	358,294	-8.63
SPRC ($\omega = 20$)	0.7	✓	337,146	-2.22
SPRC ($\omega = 0$)	0.1	✓	357,563	-8.41
SPRC ($\omega = 0$)	0.4	✓	315,182	4.44
SPRC ($\omega = 0$)	0.7	✓	288,226	12.62

Figure 7 illustrates the results reported in Tables 5, 6, and 7, which serve as a valuable reference for determining the critical CV penetration rate required for real-world implementation of the CVASC framework. Crucially, none of the proposed control schemes require on-road detectors, whereas the VAC scheme is detector-dependent. The results clearly demonstrate that the proposed control schemes generally performed similarly to or outperformed the VAC scheme. In particular, the SPRC scheme with $\omega = 0$, which fully optimizes for traffic efficiency, accounted for uncertain vehicle arrival rate due to CV penetration rate uncertainty and thus demonstrated marked superiority and robustness. Therefore, the SPRC scheme with $\omega = 0$ yielded consistent and significant improvements across most cases. At low and medium V/C ratios (i.e., 0.3 and 0.5), it outperformed the VAC scheme, even with a low CV penetration rate of 0.1, which has already been achieved in some cities around the world. At a high V/C ratio (0.7), the critical CV penetration rate was only approximately 0.3. These findings suggest that implementing the SPRC scheme with $\omega = 0$ at low CV penetration rates (i.e., 0.1–0.3) would yield better outcomes than implementing the VAC scheme across various traffic demand conditions and without requiring complete traffic information from fixed detectors. Conversely, the SPRC scheme with $\omega = 20$, which is designed to balance efficiency with stability, exhibited a slight reduction in efficiency compared with the SPRC scheme with $\omega = 0$. **Appendix Q** provides further experiments comparing the properties of the SPRC schemes with $\omega = 0$ and $\omega = 20$.

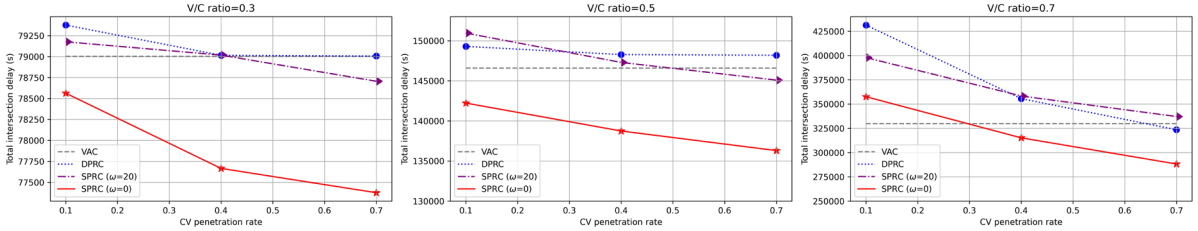


Figure 7. Comparison of different signal control schemes with varying V/C ratios and $\mathscr{p}$.

The average computational costs per cycle for DPRC, SPRC with $\omega = 0$, and SPRC with $\omega = 20$ across various V/C ratios and CV penetration rates included parameter estimations and optimization processes. These computations took approximately 0.80, 1.05, and 1.06 s, respectively, on the designated machine. These values were all below the reserved computation time of 3 s. Furthermore, more powerful graphical processing units and clusters hold the potential to further reduce computing time in practical applications. These results consistently demonstrate the applicability and efficiency of the proposed DPRC, SPRC and solution methods, highlighting the significance of incorporating the CV penetration rate uncertainty, and thus the uncertainty in vehicle arrival rate, into signal optimization.

7. Conclusions

This paper proposes a CVASC framework that adaptively optimizes signal timings cycle by cycle using CV data. Moreover, an efficient ASD model and a G^3S^2 algorithm are introduced to solve the proposed DPRC and SPRC in CVASC, and the significance of considering CV penetration rate uncertainty in adaptive signal control is emphasized. A key challenge in this study lies in accurately estimating the mean, variance, and gradients of the stochastic objective, given numerous uncertain parameters and control variables. This paper effectively breaks down the complex problem according to the first principle. The explicit modeling of the uncertainty and its propagation process plays a vital role in solving the above stochastic problem. Conducting high-dimensional and non-convex optimization is another challenge. Starting with different initial points is important for the solution algorithm to search for better local minima. The basic idea is to systematically sample the initial solution from a feasible region as large as possible, thereby increasing the likelihood of identifying superior solutions. However, implementing such a simple idea in a high-dimensional optimization problem with numerous constraints is not straightforward. Transforming this complex problem into a standard LP problem effectively resolves this dilemma. Extensive numerical experiments and VISSIM simulations demonstrate the applicability and effectiveness of the proposed models. For users prioritizing efficiency, SPRC with $\omega = 0$ is recommended, while SPRC with $\omega > 0$ is recommended to achieve a balance between efficiency and stability.

Nonetheless, this study presents certain limitations: (1) the successor matrix remains constant over cycles, and (2) complex interactions between intersections are not yet modeled. The successor matrix could be treated as variables and jointly optimized in each cycle. Furthermore, taking into account interactions between junctions delves into the more challenging domain of network-wide traffic control. Future research directions may involve addressing these limitations and extending these methods to network-wide adaptive signal control systems.

Acknowledgments

The first author was supported by a Postgraduate Scholarship from The University of Hong Kong. The second author was supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (Project Nos.: 17204919 and 17205822), and Francis S Y Bong Professorship in Engineering.

References

Al Islam, S. B., Hajbabaie, A., Aziz, H. A., 2020. A real-time network-level traffic signal control methodology with partial connected vehicle information. *Transportation Research Part C: Emerging Technologies* 121, 102830.

Ambühl, L., Menendez, M., 2016. Data fusion algorithm for macroscopic fundamental diagram estimation. *Transportation Research Part C: Emerging Technologies* 71, 184-197.

Cao, Y., Tang, K., Sun, J., Ji, Y., 2021. Day-to-day dynamic origin–destination flow estimation using connected vehicle trajectories and automatic vehicle identification data. *Transportation Research Part C: Emerging Technologies* 129, 103241.

Comert, G., 2013. Simple analytical models for estimating the queue lengths from probe vehicles at traffic signals. *Transportation Research Part B: Methodological* 55, 59-74.

Comert, G., 2016. Queue length estimation from probe vehicles at isolated intersections: Estimators for primary parameters. *European Journal of Operational Research* 252, 502-521.

Comert, G., Cetin, M., 2009. Queue length estimation from probe vehicle location and the impacts of sample size. *European Journal of Operational Research* 197, 196-202.

Comert, G., Cetin, M., 2011. Analytical evaluation of the error in queue length estimation at traffic signals from prove vehicle data. *IEEE Transactions on Intelligent Transportation Systems* 12(2), 563-573.

Du, J., Rakha, H., Gayah, V.V., 2016. Deriving macroscopic fundamental diagrams from probe data: Issues and proposed solutions. *Transportation Research Part C: Emerging Technologies* 66, 136-149.

Federal Highway Administration, 2006. Next generation simulation: Peachtree Street dataset. Accessed June 25, 2022, https://data.transportation.gov/Automobiles/Next-Generation-Simulation-NGSIM-Program-Peachtree/mupt-aksf.

Feng, Y., Head, K.L., Khoshmagham, S., Zamanipour, M., 2015. A real-time adaptive signal control in a connected vehicle environment. *Transportation Research Part C: Emerging Technologies* 55, 460-473.

Feng, Y., Zheng, J., Liu, H. X., 2018. Real-time detector-free adaptive signal control with low penetration of connected vehicles. *Transportation Research Record* 2672(18), 35-44.

Geroliminis, N., Daganzo, C.F., 2008. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological* 42(9), 759-770.

Hao, P., Ban, X.J., Guo, D., Ji, Q., 2014. Cycle-by-cycle intersection queue length distribution estimation using sample travel times. *Transportation Research Part B: Methodological* 68, 185-204.

Iqbal, M.S., Hadi, M., Xiao, Y., 2018. Effect of link-level variations of connected vehicles (CV) proportions on the accuracy and reliability of travel time estimation. *IEEE Transactions on Intelligent Transportation Systems* 20(1), 87-96.

Jenelius, E., Koutsopoulos, H.N., 2013. Travel time estimation for urban road networks using low frequency probe vehicle data. *Transportation Research Part B: Methodological* 53, 64-81.

Jenelius, E., Koutsopoulos, H.N., 2015. Probe vehicle data sampled by time or space: Consistent travel time allocation and estimation. *Transportation Research Part B: Methodological* 71, 120-137.

Jia, S., Wong, S.C., Wong, W., 2023. Uncertainty estimation of connected vehicle penetration rate. *Transportation Science* 57(5),1160-1176.

Jia, S., Wong, S.C., Wong, W., 2024a. Modeling residual-vehicle effects on uncertainty estimation of the connected vehicle penetration rate. *Transportation Research Part C: Emerging Technologies* 168, 104825.

Jia, S., Wong, S.C., Wong, W., 2024b. Modeling residual-vehicle effects in undersaturation conditions on uncertainty estimation of the connected vehicle penetration rate. *In the 25th International Symposium on Transportation and Traffic Theory*, 15-17 July, Michigan, USA.

Jia, S., Wong, S.C., Wong, W., 2024c. Estimating real-time traffic state of holding vehicles at signalized intersections using partial connected vehicle trajectory data. *Transportation Research Part C: Emerging Technologies*. Under review.

Khan, S.M., Dey, K.C., Chowdhury, M., 2017. Real-time traffic state estimation with connected vehicles. *IEEE Transactions on Intelligent Transportation Systems* 18(7), 1687-1699.

Lee, S., Wong, S.C. and Varaiya, P., 2017a. Group-based hierarchical adaptive traffic-signal control part I: Formulation. *Transportation Research Part B: Methodological* 105, 1-18.

Lee, S., Wong, S.C. and Varaiya, P., 2017b. Group-based hierarchical adaptive traffic-signal control Part II: Implementation. *Transportation Research Part B: Methodological* 104, 376-397.

Liang, X. J., Guler, S. I., Gayah, V. V., 2020a. An equitable traffic signal control scheme at isolated signalized intersections using connected vehicle technology. *Transportation Research Part C: Emerging Technologies* 110, 81-97.

Liang, X., Guler, S.I., Gayah, V.V., 2020b. Traffic signal control optimization in a connected vehicle environment considering pedestrians. *Transportation Research Record* 2674(10), 499-511.

Meng, F., Wong, S.C., Wong, W., Li, Y.C., 2017a. Estimation of scaling factors for traffic counts based on stationary and mobile sources of data. *International Journal of Intelligent Transportation Systems Research* 15(3), 180-191.

Mo, Z., Li, W., Fu, Y., Ruan, K., Di, X., 2022. CVLight: Decentralized learning for adaptive traffic signal control with connected vehicles. *Transportation Research Part C: Emerging Technologies* 141, 103728.

Rafter, C. B., Anvari, B., Box, S., Cherrett, T., 2020. Augmenting traffic signal control systems for urban road networks with connected vehicles. *IEEE Transactions on Intelligent Transportation Systems* 21(4), 1728-1740.

Rahmani, M., Jenelius, E., Koutsopoulos, H.N., 2015. Non-parametric estimation of route travel time distributions from low-frequency floating car data. *Transportation Research Part C: Emerging Technologies* 58, 343-362.

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

Tan, C., Yang, K., 2024. Privacy-preserving adaptive traffic signal control in a connected vehicle environment. *Transportation research part C: Emerging Technologies* 158, 104453.

Wang, P., Zhang, J., Deng, H., Zhang, M., 2020. Real-time urban regional route planning model for connected vehicles based on V2X communication. *Journal of Transport and Land Use* 13(1), 517-538.

Wang, Q., Yuan, Y., Yang, X. T., Huang, Z., 2021. Adaptive and multi-path progression signal control under connected vehicle environment. *Transportation Research Part C: Emerging Technologies* 124, 102965.

Wong, S.C., 1996. Group-based optimisation of signal timings using the TRANSYT traffic model. *Transportation Research Part B: Methodological* 30(3), 217-244.

Wong, W., Wong, S.C., 2015. Systematic bias in transport model calibration arising from the variability of linear data projection. *Transportation Research Part B: Methodological* 75, 1-18.

Wong, W., Wong, S.C., 2016a. Biased standard error estimations in transport model calibration due to heteroscedasticity arising from the variability of linear data projection. *Transportation Research Part B: Methodological* 88, 72-92.

Wong, W., Wong S.C., 2016b. Evaluation of the impact of traffic incidents using GPS data. *Proceedings of the Institution of Civil Engineers – Transport* 169(3), 148-162.

Wong, W., Wong S.C., 2016c. Network topological effects on the macroscopic Bureau of Public Roads function. *Transportmetrica A: Transport Science* 12(3), 272-296.

Wong, W., Wong, S.C., 2019. Unbiased estimation methods of nonlinear transport models based on linearly projected data. *Transportation Science* 53(3), 665-682.

Wong, W., Wong, S.C., Liu, X., 2019. Bootstrap standard error estimations of nonlinear transport models based on linearly projected data. *Transportmetrica A: Transport Science* 15(2), 602-630.

Wong, W., Shen, S, Zhao, Y., Liu, X., 2019. On the estimation of connected vehicle penetration rate based on single-source connected vehicle data. *Transportation Research Part B: Methodological* 126, 169-191.

Wong, W., Wong, S.C., Liu, X., 2021. Network topological effects on the macroscopic fundamental diagram. *Transportmetrica B: Transport Dynamics* 9(1), 376-398.

Xiu, D., 2010. *Numerical methods for stochastic computations: a spectral method approach*. New Jersey: Princeton University Press.

Yang, X., Lu, Y., Hao, W., 2017. Origin-destination estimation using probe vehicle trajectory and link counts. *Journal of Advanced Transportation* 2017, 4341532.

Yin, B., Menendez, M., Yang, K., 2021. Joint optimization of intersection control and trajectory planning accounting for pedestrians in a connected and automated vehicle environment. *Sustainability* 13(3), 1135.

Yin, Y., 2008. Robust optimal traffic signal timing. *Transportation Research Part B: Methodological* 42(10), 911-924.

Zhao, Y., Zheng, J., Wong, W., Wang, X., Meng, Y., Liu, H.X., 2019a. Estimation of queue lengths, probe vehicle penetration rates, and traffic volumes at signalized intersections using probe vehicle trajectories. *Transportation Research Record* 2673(11), 660-670.

Zhao, Y., Zheng, J., Wong, W., Wang, X., Meng, Y., Liu, H.X., 2019b. Various methods for queue length and traffic volume estimation using probe vehicle trajectories. *Transportation Research Part C: Emerging Technologies* 107, 70-91.

Zhao, Y., Wong, W., Zheng, J., Liu, H.X., 2022. Maximum likelihood estimation of probe vehicle penetration rates and queue length distributions from probe vehicle data. *IEEE Transactions on Intelligent Transportation Systems* 23(7), 7628-7636.

Appendix A. Glossary and table of symbols

This appendix provides a glossary (Table A1) and table of symbols (Table A2).

Table A1. Glossary.

Terminology	Description
Connected vehicles (CVs)	Vehicles equipped to exchange real-time traffic information (e.g., time, speed, location) with one another.
CV-based adaptive signal control (CVASC)	A framework that optimizes signal plans for isolation intersections on a cycle-by-cycle basis, exploiting partial CV data and operating without on-road detectors.
Consequential system delay (CSD)	An analytical model for estimating total junction delay.
Deterministic penetration rate control (DPRC)	A CV-based adaptive signal control method for isolated intersections that does not incorporate CV penetration rate uncertainty.
Stochastic penetration rate control (SPRC)	A CV-based adaptive signal control method for isolated intersections that incorporates CV penetration rate uncertainty.
Gradient-guided golden section search (G^3S^2) algorithm	An optimization algorithm that combines gradient information and golden section search for solving DPRC and SPRC.
Analytical stochastic delay (ASD) model	An analytical model used for estimating stochastic delay and its gradient.
Generalized polynomial chaos expansion (gPCE)	A polynomial approximation method used in stochastic modeling.
VISSIM	Verkehr In Städten – SIMulationsmodell, a microscopic, multi-modal traffic flow simulation software developed by PTV Planung Transport Verkehr AG in Karlsruhe, Germany.
Next-generation simulation (NGSIM) dataset	A real-world trajectory dataset collected in the United States.
Fifth-generation (5G)	An advanced communication technology supporting high-speed data transmission.
CV penetration rate	The ratio of the number of CVs to the total number of vehicles.
Single-source data penetration rate estimator (SSDPRE)	An analytical model for estimating CV penetration rate.
Probabilistic penetration rate (PPR)	An analytical model for estimating CV penetration rate uncertainty.

Markov-constrained queue length (MCQL)	An analytical model to account for residual-vehicle effects on CV penetration rate uncertainty.
Successor matrix	A matrix specifying the sequence in which traffic signal groups operate.
Clearance time matrix	A matrix defining the minimum gaps required for transitions between incompatible traffic signal groups.
Holding vehicles	Vehicles that remain undischarged at the end of an arbitrarily defined cycle.
Consequential delay	The extra delay experienced beyond regular delay due to holding vehicles.
Monte Carlo sampling (MCS)	A computational technique for generating samples from a specific probability distribution.
Vehicle-actuated control (VAC)	An adaptive traffic signal control method that dynamically adjusts signals based on real-time vehicle presence detected by on-road sensors.

Table A2. Symbols and definitions.

Symbol	Description
l	Total number of approaching lanes at an intersection.
N_T	Number of traffic signal groups.
N_P	Number of pedestrian groups.
$\mathbf{\Omega}$	Successor matrix defining signal group sequences.
Ω_{ij}	Element located in the i^{th} row and the j^{th} column in the successor matrix $\mathbf{\Omega}$.
$\mathbf{T}$	Clearance time matrix, which indicates minimum transition gaps between incompatible traffic signal groups.
t_{ij}	Element located in the i^{th} row and the j^{th} column in the clearance time matrix $\mathbf{T}$.
$\mathbf{\Psi}$	Set of incompatible signal group pairs.
D	Total junction delay.
$E(D)$	Expected total junction delay.
$Var(D)$	Variance of total junction delay.
$q_{m,k+1}$	Vehicle arrival rate in lane m in cycle $k+1$.
$n_{m,k}$	Number of CVs in lane m in cycle k.
C_k	Length of cycle k.
$\bar{q}_{m,k}$	Average arrival rate in lane m, estimated at the end of cycle k.

$p_{m,k}$	CV penetration rate in lane m in cycle k.
$Var(p_{m,k})$	Variance of $p_{m,k}$.
θ_i	Time from the cycle origin to the start of an actual green signal for control group i divided by the cycle time.
ϕ_i	Duration of the actual green signal for control group i divided by the cycle time.
ζ	Reciprocal of the cycle length.
$R_{m,k}$	Number of holding vehicles in lane m in cycle k.
$d_{m,k+1}$	Traffic delay for lane m in cycle $k+1$.
$d^c_{m,k+1}$	Consequential delay for lane m in cycle $k+1$.
f	Abstract function representing total junction delay.
C_{max}	Maximum cycle length.
g_T	Minimum traffic signal group duration.
g_P	Minimum pedestrian signal group duration.
J_D	Deterministic total junction delay.
$\mathbb{V}$	Control variable set.
J_S	Stochastic total junction delay.
ω	User-defined parameter representing the trade-off between efficiency and stability.
α_m	Probability that a vehicle selects lane m.
Φ	Orthonormal polynomials.
$\boldsymbol{c}$	Coefficient vector associated with orthonormal polynomials Φ.
K	Maximum order of generalized polynomial chaos expansion.
ξ	Number of samples for CV penetration rate.
Z_m	Random variable associated with $p_{m,k}$.
$\mathbb{V}^{(k,*)}$	Optimized signal plan for cycle k.
$\mathbb{p}$	Golden section number.
$\mathscr{p}$	True CV penetration rate.

Appendix B. Estimations of essential inputs

B.1. Estimation of CV penetration rate uncertainty

Consider any lane connected to a signalized intersection, where vehicles are required to stop during red signals and form constrained queues (Wong et al., 2019; Jia et al., 2023, 2024a, 2024b). Let $n_{m,k}$ be the number of observed CVs in the constrained queue in lane m in cycle k and $\widetilde{N}_{m,k}$ be the number of vehicles prior to the last stopped CV in the constrained queue in lane m in cycle k. According

to these two quantities, Wong et al. (2019) proved that the SSDPRE, $E(\tilde{p}_{m,k})$, is an unbiased estimator for the average CV penetration rate, $\bar{p}_{m,k}$, where

$$\tilde{p}_{m,k} = \begin{cases} \dfrac{n_{m,k}-1}{\widetilde{N}_{m,k}-1} & if\ n_{m,k} > 1\ and\ \widetilde{N}_{m,k} > 1 \\ 1 & if\ n_{m,k} = 1\ and\ \widetilde{N}_{m,k} = 1 \\ 0 & if\ n_{m,k} = 1\ and\ \widetilde{N}_{m,k} > 1 \\ 0 & if\ n_{m,k} = 0\ and\ \widetilde{N}_{m,k} = 0 \end{cases}, \forall m \in [1, l]. \quad \text{(B1)}$$

More details can be found in Wong et al. (2019). However, in real-time applications, only a single realization of $\tilde{p}_{m,k}$ is available, and this can deviate from $\bar{p}_{m,k}$. To ensure unbiased model estimations and optimal system optimizations, it is crucial to account for the uncertainty of $\tilde{p}_{m,k}$, which is measured by the variance of the distribution of $\tilde{p}_{m,k}$, $Var(\tilde{p}_{m,k})$. Considering any constrained queue length $N_{m,k}$ that follows a counting distribution such that $P(N_{m,k} = i) = \xi_i, \forall i = 0, 1, 2, \ldots$ and $n_{m,k}$, which follows a binomial distribution, $B(N_{m,k}, \bar{p}_{m,k})$, Jia et al. (2023) derived an exact PPR model quantifying $Var(\tilde{p}_{m,k})$:

$$Var(\tilde{p}_{m,k}) = \lim_{j \to +\infty} \left[\sum_{i=1}^{j} \xi_i\, V_2(i, \bar{p}_{m,k})\right], \forall m \in [1, l], \quad \text{(B2)}$$

where

$$V_2(N_{m,k}, \bar{p}_{m,k}) = \begin{cases} \sum_{i=2}^{N_{m,k}} {\bar{p}_{m,k}}^{i} (1-\bar{p}_{m,k})^{N_{m,k}-i} \left[V_1(i, N_{m,k}) + \left(\frac{i}{N_{m,k}}\right)^2\right] \binom{N_{m,k}}{i} - {\bar{p}_{m,k}}^2 + \bar{p}_{m,k}(1-\bar{p}_{m,k})^{N_{m,k}-1} & if\ N_{m,k} > 1 \\ \bar{p}_{m,k}(1-\bar{p}_{m,k}) & if\ N_{m,k} = 1 \end{cases}, \quad \text{(B3)}$$

and

$$V_1(n_{m,k}, N_{m,k}) = \begin{cases} \dfrac{\sum_{i=1}^{N_{m,k}-n_{m,k}+1} \frac{n_{m,k}-1}{N_{m,k}-i} \binom{N_{m,k}-i-1}{n_{m,k}-2}}{\binom{N}{n}} - \dfrac{n^2}{N^2} & if\ n_{m,k} > 1 \\ \dfrac{{n_{m,k}}^2 - 2n_{m,k} + N_{m,k}}{{N_{m,k}}^2} & if\ n_{m,k} = 1 \\ 0 & if\ n_{m,k} = 0 \end{cases}. \quad \text{(B4)}$$

Thus, $Var(\tilde{p}_{m,k})$ can be taken as the estimator for $Var(p_{m,k})$. Under the assumption that the distribution of $N_{m,k}$, which is the essential input of the PPR model, follows a Poisson distribution, the governing parameter of the constrained queue length distribution N_0 can be estimated using the constant dissipation time (CDT) model (Jia et al., 2023):

$$N_0 = \frac{s\bar{q}_{m,k} r}{s - \bar{q}_{m,k}}, \quad \text{(B5)}$$

where s and r represent the saturation flow rate and the red period, respectively. Furthermore, to account for the complex residual-vehicle effects, an MCQL model can be used to estimate the distribution of $N_{m,k}$. More details on the development of the model are presented in Jia et al. (2024a, 2024b).

B.2. Estimation of average arrival rate and CV penetration rate

Given the PPR and CDT models, the joint probability distribution of $n_{m,k}$ and $\widetilde{N}_{m,k}$ is given as follows (Jia et al., 2023):

$$P\left(n_{m,k}=i,\widetilde{N}_{m,k}=j\right)=\begin{cases}\pi_0+\sum_{z=1}^{k}\pi_z\left(1-\bar{p}_{m,k}\right)^z, & i=0,j=0\\ \sum_{z=j}^{k}\pi_z\binom{j-1}{i-1}\bar{p}_{m,k}{}^i\left(1-\bar{p}_{m,k}\right)^{z-i}, & \forall i,j=1,2,\dots,k,j\geq i\end{cases}, \quad \text{(B6)}$$

where $\pi_z=P\left(N_{m,k}=z\right),\forall z=0,1,2,\dots$. Thus, $\bar{q}_{m,k}$ and $\bar{p}_{m,k}$ can be estimated using the following maximum likelihood estimation formulation (Jia et al., 2023, 2024a, 2024b, 2024c):

$$\max_{\bar{q}_{m,k},\bar{p}_{m,k}}\prod_{j=0}^{\omega}P(n_{m,k-j},\widetilde{N}_{m,k-j}),\forall m\in[1,l], \quad \text{(B7)}$$

where $\omega=0,1,2,\dots,k-1$, and it represents the number of past cycles considered in the likelihood function. Following Jia et al. (2023, 2024a, 2024b, 2024c), this paper sets ω to 2. A simple grid search method can be used to solve Eq. (B7) and search for the optimal solution.

B.3. Estimation of the number of holding vehicles

The number of holding vehicles at the end of a cycle serves as the initial state of the subsequent cycle and plays a pivotal role in the estimation of traffic delay in the following cycle. Integrating holding vehicle information into the estimation of $Var(p_{m,k})$ and adaptive signal control can significantly enhance traffic efficiency (Jia et al., 2024a, 2024b). To accurately estimate the number of holding vehicles in lane m in cycle k, $R_{m,k}$, a generic and fully analytical model, CV-based holding vehicle (CVHV), was derived in a previous study (Jia et al., 2024c). The model consists of two sub-models, CVHV-I and CVHV-II, which are tailored to accommodate various holding vehicle patterns arising from different signal structures. Specifically, when a lane is controlled by a signal plan ending with an effective red, the CVHV-I sub-model should be applied; otherwise, the CVHV-II sub-model should be used. The estimated $R_{m,k}$ is used in traffic delay estimation, as elaborated in Section 3.

Appendix C. Proof of Proposition 1

Under the assumption of a uniform vehicle arrival pattern, the traffic delay for lane m in cycle $k+1$, $d_{m,k+1}$, can be visualized as the area enclosed by the cumulative arrival and departure lines (Figure 3). Cases 1 and 2 in Figure 3 are utilized to derive $d_{m,k+1}^{(1)}$ and $d_{m,k+1}^{(2)}$, respectively. The formulae for $d_{m,k+1}^{(1)}$ and $d_{m,k+1}^{(2)}$ are derived according to simple geometry (Tables 1 and 2). It should be noted that the estimated delay is primarily governed by the vehicle arrival rate, which dictates the slope of the cumulative arrival line and largely determines the total number of arriving vehicles in a cycle. If the CV penetration rate uncertainty is ignored, a fixed vehicle arrival rate and slope are determined, resulting in the DPRC. In contrast, incorporating CV penetration rate uncertainty accounts

for the uncertainty in vehicle arrival rate, and hence the uncertainty in the slope of the cumulative arrival line, establishing the SPRC.

Appendix D. Derivation of $d^{c}_{m,k+1}$

Under the assumption that holding vehicles arise owing to temporarily high demand in lane m in cycle $k+1$, $q_{m,k+1}$, and the arrival rate in lane m returns to the estimated average arrival rate in cycle k, $\bar{q}_{m,k}$, after cycle $k+1$, the number of holding vehicles in lane m at the end of cycle $k+1$, $R_{m,k+1}$, can be estimated as $R^{(1)}_{m,k+1}$ if the signal group controlling that lane ends with effective red; otherwise, it is estimated as $R^{(2)}_{m,k+1}$, where

$$R^{(1)}_{m,k+1} = \begin{cases} q_{m,k+1} r_2 & if\ q_{m,k+1} \le \dfrac{sg - R_{m,k}}{r_1 + g} \\ R_{m,k} + q_{m,k+1} C - sg & if\ q_{m,k+1} > \dfrac{sg - R_{m,k}}{r_1 + g} \end{cases}; \tag{D1}$$

and

$$R^{(2)}_{m,k+1} = \begin{cases} q_{m,k+1}(r + g_2) - sg_2 & if\ q_{m,k+1} \le \dfrac{sg_1 - R_{m,k}}{g_1}, q_{m,k+1} > \dfrac{sg_2}{r + g_2} \\ R_{m,k} + q_{m,k+1} C - sg & if\ q_{m,k+1} > \dfrac{sg_1 - R_{m,k}}{g_1}, q_{m,k+1} > \dfrac{sg - R_{m,k}}{C}. \\ 0 & otherwise \end{cases} \tag{D2}$$

For ease of analysis, $R_{m,k+1}$ can be projected to the end of the nearest green period, denoted as $R_{m,k+1}{}'$, which can be estimated as either $R^{(1)}_{m,k+1}{}'$ or $R^{(2)}_{m,k+1}{}'$ depending on whether the lane is controlled by a signal group ending with effective red or green, as shown in Tables 3 and 4.

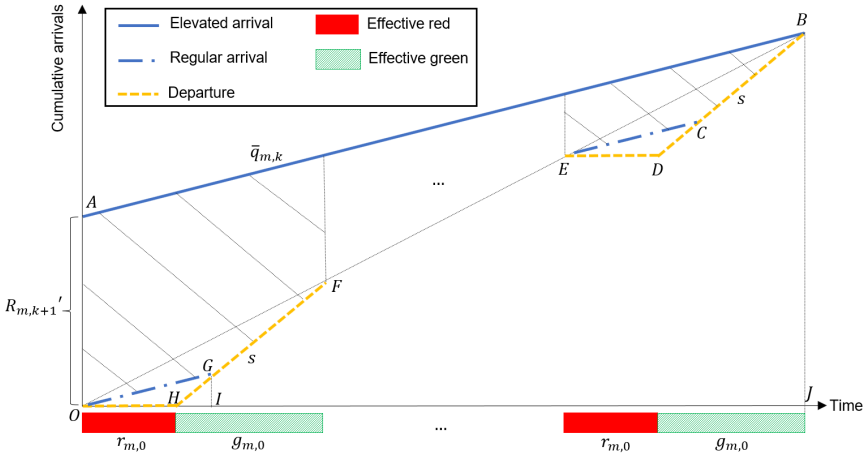


Figure D1. Illustration of consequential delay estimation.

An elevated arrival line is defined as a heightened level of arrival due to the presence of holding vehicles, while a regular arrival line represents the anticipated level of arrival without holding vehicles. As depicted in Figure D1, $R_{m,k+1}{}' > 0$ indicates that some vehicles are carried over to the subsequent cycle owing to the temporarily high demand. In such cases, the new arrivals in the following cycle would experience extra delays beyond the regular delays experienced at $R_{m,k+1}{}' = 0$ owing to the first-

come-first-served property. Considering a selected analysis signal plan with an analysis cycle length C_0, analysis effective green $g_{m,0}$, and analysis effective red $r_{m,0}$ as a reference, the total delay for the system with holding vehicles is determined by the area enclosed by the elevated arrival line and the departure line. Conversely, the regular delay for the system without holding vehicles is the area enclosed by the regular arrival line and the departure line. The consequential delay induced by ${R_{m,k+1}}'$ as expressed in Eq. (11) is graphically represented by the difference between the total delay and the regular delay (Figure D1). In contrast, ${R_{m,k+1}}' = 0$ indicates that all of the arrived vehicles in cycle $k+1$ are discharged before the end of the green period. Consequently, no vehicles will be carried over to the subsequent cycles, resulting in no additional influence on the system performance. Therefore, the consequential delay, $d_{m,k+1}^c$, is expected to be 0.

According to Figure D1,

$$d_{m,k+1}^c = S_{OAB} + \frac{|OJ|}{C_0} S_{OFG}, \tag{D3}$$

where

$$S_{OAB} = \frac{1}{2} {R_{m,k+1}}' |OJ|, \tag{D4}$$

$$S_{OFG} = S_{OFH} - S_{OGH} = \frac{1}{2} s r_{m,0} g_{m,0} - \frac{1}{2} r_{m,0} s |HI|. \tag{D5}$$

The conservation law of traffic flow guarantees that

$${R_{m,k+1}}' + \bar{q}_{m,k} |OJ| = s \frac{g_{m,0}}{C_0} |OJ|, \tag{D6}$$

and

$$\bar{q}_{m,k} (r_{m,0} + |HI|) = s|HI|. \tag{D7}$$

Rearranging terms, Eqs. (D6) and (D7) respectively become

$$|OJ| = \frac{C_0 {R_{m,k+1}}'}{s g_{m,0} - C_0 \bar{q}_{m,k}}, \tag{D8}$$

and

$$|HI| = \frac{\bar{q}_{m,k} r_{m,0}}{s - \bar{q}_{m,k}}. \tag{D9}$$

By substituting Eqs. (D8) and (D9) into Eqs. (D4) and (D5), respectively, Eq. (D3) becomes

$$d_{m,k+1}^c = \gamma_{1,m} {R_{m,k+1}}'^2 + \gamma_{2,m} {R_{m,k+1}}', \forall m \in [1, l], \tag{D10}$$

where

$$\gamma_{1,m} = \frac{C_0}{2\left(s g_{m,0} - C_0 \bar{q}_{m,k}\right)}, \tag{D11}$$

$$\gamma_{2,m} = \frac{r_{m,0} s}{2(s - \bar{q}_{m,k})}. \tag{D12}$$

As the coefficients $\gamma_{1,m}$ and $\gamma_{2,m}$ in Eq. (D10) are lane-specific, the consequential delays for different lanes can differ even if the projected numbers of holding vehicles are identical. Moreover, the inverse

forms of $\gamma_{1,m}$ and $\gamma_{2,m}$ are sensitive to signal plans. Therefore, the averages of lane-specific coefficients γ_1 and γ_2, as shown in Eqs. (11) to (13), are adopted in **Proposition 2**.

Appendix E. Gradient of $d_{m,k+1}$

According to $d_{m,k+1}^{(1)}$ and $d_{m,k+1}^{(2)}$ in Tables 1 and 2 and the parameters defined in Eqs. (4)–(10), the first partial derivatives of $d_{m,k+1}^{(1)}$ and $d_{m,k+1}^{(2)}$ $w.r.t.$ different parameters are derived using the chain rules, as shown in Tables E1 and E2, respectively.

Table E1. First partial derivatives of $d_{m,k+1}^{(1)}$.

Condition	First partial derivative
$q_{m,k+1} \le \frac{sg - R_{m,k}}{r_1 + g}$	$\frac{\partial d_{m,k+1}^{(1)}}{\partial \theta_i} = \frac{R_{m,k} + q_{m,k+1}(r_1 - r_2)}{\zeta} + \frac{q_{m,k+1}(R_{m,k} + q_{m,k+1}r_1)}{\zeta(s - q_{m,k+1})}$
	$\frac{\partial d_{m,k+1}^{(1)}}{\partial \phi_i} = -\frac{q_{m,k+1}r_2}{\zeta}$
	$\frac{\partial d_{m,k+1}^{(1)}}{\partial \zeta} = -\frac{(R_{m,k} + q_{m,k+1}r_1)\theta_i + q_{m,k+1}r_2(1 - \theta_i - \phi_i)}{\zeta^2} - \frac{2q_{m,k+1}\theta_i(R_{m,k} + q_{m,k+1}r_1)}{2\zeta^2(s - q_{m,k+1})}$
$q_{m,k+1} > \frac{sg - R_{m,k}}{r_1 + g}$	$\frac{\partial d_{m,k+1}^{(1)}}{\partial \theta_i} = \frac{sg}{\zeta}$
	$\frac{\partial d_{m,k+1}^{(1)}}{\partial \phi_i} = \frac{s(g + r_1 - C)}{\zeta}$
	$\frac{\partial d_{m,k+1}^{(1)}}{\partial \zeta} = -\frac{R_{m,k} + q_{m,k+1}C - sg + \phi_i s(g + r_1 - C) + sg\theta_i}{\zeta^2}$

Table E2. First partial derivatives of $d_{m,k+1}^{(2)}$.

Condition	First partial derivative
$q_{m,k+1} \le \frac{sg_1 - R_{m,k}}{g_1}$, $q_{m,k+1} \le \frac{sg_2}{r + g_2}$	$\frac{\partial d_{m,k+1}^{(2)}}{\partial \theta_i} = 0$
	$\frac{\partial d_{m,k+1}^{(2)}}{\partial \phi_i} = \frac{sq_{m,k+1}r}{\zeta(q_{m,k+1} - s)}$
	$\frac{\partial d_{m,k+1}^{(2)}}{\partial \zeta} = -\frac{sq_{m,k+1}r(1 - \phi_i)}{\zeta^2(s - q_{m,k+1})}$
$q_{m,k+1} \le \frac{sg_1 - R_{m,k}}{g_1}$,	$\frac{\partial d_{m,k+1}^{(2)}}{\partial \theta_i} = \frac{sg_2 - q_{m,k+1}(r + g_2)}{\zeta}$

$q_{m,k+1} > \frac{sg_2}{r+g_2}$	$\frac{\partial d_{m,k+1}^{(2)}}{\partial \phi_i} = -\frac{q_{m,k+1}(r+g_2)}{\zeta}$ $\frac{\partial d_{m,k+1}^{(2)}}{\partial \zeta} = -\frac{q_{m,k+1} r(2-\phi_i-\theta_i)}{\zeta^2}$ $-\frac{g_2\left[q_{m,k+1}(1-\phi_i)+(1-\theta_i)\left(q_{m,k+1}-s\right)\right]}{\zeta^2}$
$q_{m,k+1} > \frac{sg_1 - R_{m,k}}{g_1}$, $q_{m,k+1} \le \frac{sg - R_{m,k}}{C}$	$\frac{\partial d_{m,k+1}^{(2)}}{\partial \theta_i} = -\frac{sr}{\zeta}$ $\frac{\partial d_{m,k+1}^{(2)}}{\partial \phi_i} = \frac{s\left[rs + R_{m,k} + g_1\left(q_{m,k+1}-s\right)\right]}{\zeta\left(q_{m,k+1}-s\right)}$ $\frac{\partial d_{m,k+1}^{(2)}}{\partial \zeta} = -\frac{1}{\zeta^2}\{(\theta_i+\phi_i-1)\left[R_{m,k}+\left(q_{m,k+1}-s\right)(g_1+r)\right]$ $+(1-\phi_i)\left[R_{m,k}+q_{m,k+1}r$ $+g_1\left(q_{m,k+1}-s\right)\right]\}$ $+\frac{R_{m,k}+g_1\left(q_{m,k+1}-s\right)+q_{m,k+1}r}{2\left(s-q_{m,k+1}\right)^2}\Big\{2\left(s$ $-q_{m,k+1}\right)\Big\{-\frac{1}{\zeta^2}\left[(\theta_i+\phi_i-1)\left(q_{m,k+1}-s\right)\right.$ $\left.+q_{m,k+1}(1-\phi_i)\right]\Big\}\Big\}$
$q_{m,k+1} > \frac{sg_1 - R_{m,k}}{g_1}$, $q_{m,k+1} > \frac{sg - R_{m,k}}{C}$	$\frac{\partial d_{m,k+1}^{(2)}}{\partial \theta_i} = -\frac{sr}{\zeta}$ $\frac{\partial d_{m,k+1}^{(2)}}{\partial \phi_i} = -\frac{s(r+g_2)}{\zeta}$ $\frac{\partial d_{m,k+1}^{(2)}}{\partial \zeta} = -\frac{1}{\zeta^2}\{(\theta_i+\phi_i-1)\left[R_{m,k}+\left(q_{m,k+1}-s\right)(g_1+r)\right]$ $+(1-\phi_i)\left[R_{m,k}+q_{m,k+1}r$ $+g_1\left(q_{m,k+1}-s\right)\right]\}$ $+\frac{1}{2\zeta^2}\left[s(1-\theta_i)(g+g_1)\right.$ $+g_2 s(2\phi_i+\theta_i-1)-g_2 q_{m,k+1}(1+\theta_i)$ $-q_{m,k+1}(1-\theta_i)(r+g_1+C)$ $\left.-2R_{m,k}(1-\theta_i)\right]$

Appendix F. Gradient of $d_{m,k+1}^{c}$

According to Eq. (11), the first partial derivatives of $d_{m,k+1}^{c}\ w.r.t.\ x$, $\forall x \in \{\theta_i, \phi_i, \zeta\}$, are given by

$$\frac{\partial d_{m,k+1}^{c}}{\partial x} = \left(2\gamma_1 {R_{m,k+1}}' + \gamma_2\right)\frac{\partial {R_{m,k+1}}'}{\partial x}, \forall x \in \{\theta_i, \phi_i, \zeta\}. \tag{F1}$$

As ${R_{m,k+1}}'$ can be respectively estimated as either ${R_{m,k+1}^{(1)}}'$ or ${R_{m,k+1}^{(2)}}'$ depending on whether the lane is controlled by a signal group ending with effective red or green, the first partial derivatives of ${R_{m,k+1}^{(1)}}'$ and ${R_{m,k+1}^{(2)}}'$ are respectively derived in Tables F1 and F2.

Table F1. First partial derivatives of ${R_{m,k+1}^{(1)}}'$.

Conditions	$\boldsymbol{{R_{m,k+1}^{(1)}}'}$
$q_{m,k+1} \le \frac{sg - R_{m,k}}{r_1 + g}$, $\bar{q}_{m,k} > \frac{sg - q_{m,k+1} r_2}{r_1 + g}$	$\frac{\partial {R_{m,k+1}^{(1)}}'}{\partial \theta_i} = \frac{\bar{q}_{m,k} - q_{m,k+1}}{\zeta}$ $\frac{\partial {R_{m,k+1}^{(1)}}'}{\partial \phi_i} = \frac{\bar{q}_{m,k} - s - q_{m,k+1}}{\zeta}$ $\frac{\partial {R_{m,k+1}^{(1)}}'}{\partial \zeta} = \frac{(\theta_i + \phi_i)\left(q_{m,k+1} - \bar{q}_{m,k}\right) - q_{m,k+1} + s\phi_i}{\zeta^2}$
$q_{m,k+1} > \frac{sg - R_{m,k}}{r_1 + g}$, $\bar{q}_{m,k} > \frac{2sg - R_{m,k} - q_{m,k+1} C}{r_1 + g}$	$\frac{\partial {R_{m,k+1}^{(1)}}'}{\partial \theta_i} = \frac{\bar{q}_{m,k}}{\zeta}$ $\frac{\partial {R_{m,k+1}^{(1)}}'}{\partial \phi_i} = \frac{\bar{q}_{m,k} - 2s}{\zeta}$ $\frac{\partial {R_{m,k+1}^{(1)}}'}{\partial \zeta} = \frac{2s\phi_i - q_{m,k+1} - \bar{q}_{m,k}(\theta_i + \phi_i)}{\zeta^2}$

Table F2. First partial derivatives of ${R_{m,k+1}^{(2)}}'$.

Conditions	$\boldsymbol{{R_{m,k+1}^{(2)}}'}$
$q_{m,k+1} \le \frac{sg_1 - R_{m,k}}{g_1}$, $q_{m,k+1} > \frac{sg_2}{r + g_2}$, $\bar{q}_{m,k} > \frac{sg - q_{m,k+1}(r + g_2)}{g_1}$	$\frac{\partial {R_{m,k+1}^{(2)}}'}{\partial \theta_i} = \frac{\bar{q}_{m,k} - q_{m,k+1}}{\zeta}$ $\frac{\partial {R_{m,k+1}^{(2)}}'}{\partial \phi_i} = \frac{\bar{q}_{m,k} - s - q_{m,k+1}}{\zeta}$ $\frac{\partial {R_{m,k+1}^{(2)}}'}{\partial \zeta} = \frac{q_{m,k+1}(\theta_i + \phi_i - 2) + s\phi_i + \bar{q}_{m,k}(1 - \theta_i - \phi_i)}{\zeta^2}$

$q_{m,k+1} > \frac{sg_1 - R_{m,k}}{g_1}$,	$\frac{\partial R_{m,k+1}^{(2)}}{\partial \theta_i}' = \frac{\bar{q}_{m,k} - s}{\zeta}$
$q_{m,k+1} > \frac{sg - R_{m,k}}{C}$,	$\frac{\partial R_{m,k+1}^{(2)}}{\partial \phi_i}' = \frac{\bar{q}_{m,k} - 2s}{\zeta}$
$\bar{q}_{m,k} > \frac{s(g_1 + g) - R_{m,k} - q_{m,k+1}C}{g_1}$	$\frac{\partial R_{m,k+1}^{(2)}}{\partial \zeta}' = \frac{(\bar{q}_{m,k} - s)(1 - \theta_i - \phi_i) + s\phi_i - q_{m,k+1}}{\zeta^2}$

Appendix G. Proof of Lemma 1

Let q_m represent the traffic demand for lane $m, \forall m \in [1, l]$, and let q denote the total traffic demand, where $q = \sum_{m=1}^{l} q_m$ follows $Pois(\lambda)$. The proof is divided into two parts: (1) the independence of q_m, $\forall m \in [1, l]$, and (2) the independence of $p_{m,k}$, $\forall m \in [1, l]$.

(1) Independence of q_m

Let the traffic demand assigned to lane m be $v_m, \forall m \in [1, l]$. The probability of observing $q_1 = v_1, \dots, q_m = v_m, \dots, q_l = v_l$ is given by

$$\begin{aligned} &P(q_1 = v_1, \dots, q_m = v_m, \dots, q_l = v_l) \\ &= P\left(q_1 = v_1, \dots, q_m = v_m, \dots, q_l = v_l \middle| \sum_{m=1}^{l} q_m = \sum_{m=1}^{l} v_m\right) P\left(\sum_{m=1}^{l} q_m = \sum_{m=1}^{l} v_m\right) \\ &= P\left(q_1 = v_1, \dots, q_m = v_m, \dots, q_l = v_l \middle| q = \sum_{m=1}^{l} v_m\right) P\left(q = \sum_{m=1}^{l} v_m\right), \end{aligned} \tag{G1}$$

where $P\left(q_1 = v_1, \dots, q_m = v_m, \dots, q_l = v_l \middle| q = \sum_{m=1}^{l} v_m\right)$ is the probability of observing $q_1 = v_1, \dots, q_m = v_m, \dots, q_l = v_l$ under the condition that $q = \sum_{m=1}^{l} v_m$. Given the lane choice probability $\alpha_m, \forall m \in [1, l]$, the lane choice process can be modeled using a multinomial distribution with the following probability mass function:

$$P(q_1 = x_1, \dots, q_m = x_m, \dots, q_l = x_l) = \begin{cases} \frac{q!}{\prod_{m=1}^{l} x_m!} \prod_{m=1}^{l} {\alpha_m}^{x_m} & if\ q \in \mathbb{N}. \\ 0 & otherwise \end{cases} \tag{G2}$$

Thus,

$$P\left(q_1 = v_1, \dots, q_m = v_m, \dots, q_l = v_l \middle| q = \sum_{m=1}^{l} v_m\right) = \frac{\left(\sum_{m=1}^{l} v_m\right)!}{\prod_{m=1}^{l} v_m!} \prod_{m=1}^{l} {\alpha_m}^{v_m}. \tag{G3}$$

Moreover,

$$P\left(q = \sum_{m=1}^{l} v_m\right) = e^{-\lambda} \frac{\lambda^{\sum_{m=1}^{l} v_m}}{\left(\sum_{m=1}^{l} v_m\right)!}. \tag{G4}$$

Substituting Eqs. (G3) and (G4) into Eq. (G1) gives

$$P(q_1 = v_1, \dots, q_m = v_m, \dots, q_l = v_l) = \frac{e^{-\lambda} \prod_{m=1}^{l} {\alpha_m}^{v_m} \lambda^{\sum_{m=1}^{l} v_m}}{\prod_{m=1}^{l} v_m!} = \frac{e^{-\lambda} \prod_{m=1}^{l} (\lambda\alpha_m)^{v_m}}{\prod_{m=1}^{l} v_m!}. \quad \text{(G5)}$$

Considering that $\sum_{m=1}^{l} \alpha_m = 1$, Eq. (G5) can be rewritten as

$$P(q_1 = v_1, \dots, q_m = v_m, \dots, q_l = v_l) = \frac{e^{-\lambda \sum_{m=1}^{l} \alpha_m} \prod_{m=1}^{l} (\lambda\alpha_m)^{v_m}}{\prod_{m=1}^{l} v_m!} = \prod_{m=1}^{l} \frac{e^{-\alpha_m\lambda} (\alpha_m\lambda)^{v_m}}{v_m!}. \quad \text{(G6)}$$

In addition, as only α_m of the total demand is distributed to lane m, the traffic demand for lane m is given by $\alpha_m\lambda$. Therefore,

$$P(q_m = v_m) = \frac{e^{-\alpha_m\lambda} (\alpha_m\lambda)^{v_m}}{v_m!}, \forall m \in [1, l]. \quad \text{(G7)}$$

Substituting Eq. (G7) into Eq. (G6) gives

$$P(q_1 = v_1, \dots, q_m = v_m, \dots, q_l = v_l) = \prod_{m=1}^{l} P(q_m = v_m). \quad \text{(G8)}$$

Therefore, $q_m, \forall m \in [1, l]$ are independent of each other.

(2) Independence of $p_{m,k}$

The estimates from the SSDPRE method, $p_{m,k}$, depend on the number of CVs, $n_{m,k}$, and the total number of observable vehicles ($\widetilde{N}_{m,k}$; i.e., the number of vehicles prior to the last stopped CV, including itself) that are stopped by red signals. Considering that $\widetilde{N}_{m,k}$ includes $n_{m,k}$, they are directly correlated. As $n_{m,k}$ is dependent on q_m,

$$p_{m,k} = g_i(q_m), m \in [1, l], \quad \text{(G9)}$$

where $g_i(\cdot)$, $\forall i \in [1, l]$, represents the i^{th} real-valued mapping function. Let $p_{m,k}^*$ and ${g_i}^{-1}(p_{m,k}^*)$ be a realization of $p_{m,k}$ and the preimage of $p_{m,k}^*$ under g_i, respectively. It follows that

$$\begin{aligned} &P\left(p_{1,k} = p_{1,k}^*, \dots, p_{m,k} = p_{m,k}^*, \dots, p_{l,k} = p_{l,k}^*\right) \\ &\quad = P\left(q_1 = {g_1}^{-1}\left(p_{1,k}^*\right), \dots, q_m = {g_m}^{-1}\left(p_{m,k}^*\right), \dots, q_l = {g_l}^{-1}\left(p_{l,k}^*\right)\right). \end{aligned} \quad \text{(G10)}$$

According to the proven independence of q_m, $\forall m \in [1, l]$,

$$\begin{aligned} P\left(p_{1,k} = p_{1,k}^*, \dots, p_{m,k} = p_{m,k}^*, \dots, p_{l,k} = p_{l,k}^*\right) &= \prod_{m=1}^{l} P\left(q_m = {g_m}^{-1}\left(p_{m,k}^*\right)\right) \\ &= \prod_{m=1}^{l} P\left(p_{m,k} = p_{m,k}^*\right). \end{aligned} \quad \text{(G11)}$$

Therefore, $p_{m,k}$, $\forall m \in [1, l]$ are independent of each other.

Appendix H. Validation of Lemma 1

Let $\rho(\cdot,\cdot)$ and $q_m, \forall m \in [1,2]$, represent the correlation coefficient between two random variables and the traffic demand in lane m, respectively. To examine the independence of $p_{m,k}$, various combinations of λ, α_1, and p under groups I, II, and III are considered. The correlation coefficients

evaluated for these combinations are presented in Table H1. The near-zero correlation coefficients across different combinations of λ, α_1, and p validate **Lemma 1**.

Table H1. Correlation coefficients of traffic demands and CV penetration rates.

No.	λ	α_1	p	$\rho(q_1, q_2)$	$\rho(p_{1,k}, p_{2,k})$
Baseline	20	0.4	0.4	0.002	−0.008
I-1	10	0.4	0.4	0.006	−0.005
I-2	30	0.4	0.4	−0.008	−0.003
II-1	20	0.1	0.4	0.012	−0.008
II-2	20	0.7	0.4	−0.008	0.011
II-3	20	0.9	0.4	−0.009	0.007
III-1	20	0.4	0.1	0.002	0.002
III-2	20	0.4	0.7	0.002	0.004
III-3	20	0.4	0.9	0.002	-0.007

Appendix I. Proof of Proposition 3

Given the distribution function of $\boldsymbol{Z}$, $F_{\boldsymbol{Z}}$,

$$E(\widetilde{D}) = E_{F_{\boldsymbol{Z}}}\left(\sum_{|\boldsymbol{k}|\le K} c_{\boldsymbol{k}}\,\Phi_{\boldsymbol{k}}(\boldsymbol{Z})\right) = \sum_{|\boldsymbol{k}|\le K} c_{\boldsymbol{k}}\,E_{F_{\boldsymbol{Z}}}[\Phi_{\boldsymbol{k}}(\boldsymbol{Z})]. \tag{I1}$$

By definition, $\Phi_{\boldsymbol{0}}(\boldsymbol{Z}) = 1$. Thus, Eq. (I1) can be rewritten as

$$E(\widetilde{D}) = \sum_{|\boldsymbol{k}|\le K} c_{\boldsymbol{k}}\,E_{F_{\boldsymbol{Z}}}[\Phi_{\boldsymbol{k}}(\boldsymbol{Z})\Phi_{\boldsymbol{0}}(\boldsymbol{Z})]. \tag{I2}$$

The orthonormality of $\Phi_{\boldsymbol{k}}(\boldsymbol{Z})$, $\forall|\boldsymbol{k}| \le K$, gives

$$E_{F_{\boldsymbol{Z}}}\left[\Phi_{\boldsymbol{k}_1}(\boldsymbol{Z})\Phi_{\boldsymbol{k}_2}(\boldsymbol{Z})\right] = \begin{cases} 1 & if\ \boldsymbol{k}_1 = \boldsymbol{k}_2 \\ 0 & otherwise \end{cases}. \tag{I3}$$

Therefore,

$$E(\widetilde{D}) = c_{\boldsymbol{0}}E_{F_{\boldsymbol{Z}}}[\Phi_{\boldsymbol{0}}(\boldsymbol{Z})\Phi_{\boldsymbol{0}}(\boldsymbol{Z})] + \sum_{0<|\boldsymbol{k}|\le \mathrm{K}} c_{\boldsymbol{k}}\,E_{F_{\boldsymbol{Z}}}[\Phi_{\boldsymbol{k}}(\boldsymbol{Z})\Phi_{\boldsymbol{0}}(\boldsymbol{Z})] = c_{\boldsymbol{0}}. \tag{I4}$$

For $Var(\widetilde{D})$,

$$\begin{aligned} Var(\widetilde{D}) &= E_{F_{\boldsymbol{Z}}}\left\{\left[\widetilde{D} - E(\widetilde{D})\right]^2\right\} = E_{F_{\boldsymbol{Z}}}\left\{\left[\sum_{|\boldsymbol{k}|\le K} c_{\boldsymbol{k}}\,\Phi_{\boldsymbol{k}}(\boldsymbol{Z}) - c_{\boldsymbol{0}}\right]^2\right\} \\ &= E_{F_{\boldsymbol{Z}}}\left\{\left[\sum_{|\boldsymbol{k}|\le K} c_{\boldsymbol{k}}\,\Phi_{\boldsymbol{k}}(\boldsymbol{Z})\right]^2\right\} - 2E_{F_{\boldsymbol{Z}}}\left[c_{\boldsymbol{0}}\sum_{|\boldsymbol{k}|\le K} c_{\boldsymbol{k}}\,\Phi_{\boldsymbol{k}}(\boldsymbol{Z})\right] + E_{F_{\boldsymbol{Z}}}(c_{\boldsymbol{0}}^2). \end{aligned} \tag{I5}$$

Let $A = E_{F_{\boldsymbol{Z}}}\left\{\left[\sum_{|\boldsymbol{k}|\le K} c_{\boldsymbol{k}}\,\Phi_{\boldsymbol{k}}(\boldsymbol{Z})\right]^2\right\}$, $B = E_{F_{\boldsymbol{Z}}}\left[c_{\boldsymbol{0}}\sum_{|\boldsymbol{k}|\le K} c_{\boldsymbol{k}}\,\Phi_{\boldsymbol{k}}(\boldsymbol{Z})\right]$, and $F = E_{F_{\boldsymbol{Z}}}(c_{\boldsymbol{0}}^2)$. The total number of cases for $|\boldsymbol{k}| \le K$ is denoted as $\mathbb{Q}$. For A, it follows that

$$
\begin{aligned}
A &= E_{F_{\boldsymbol{Z}}}\left\{[c_{\boldsymbol{k}_1}\Phi_{\boldsymbol{k}_1}(\boldsymbol{Z}) + c_{\boldsymbol{k}_2}\Phi_{\boldsymbol{k}_2}(\boldsymbol{Z}) + \cdots + c_{\boldsymbol{k}_{\mathbb{Q}}}\Phi_{\boldsymbol{k}_{\mathbb{Q}}}(\boldsymbol{Z})]^2\right\} \\
&= E_{F_{\boldsymbol{Z}}}\left[c_{\boldsymbol{k}_1}^2\Phi_{\boldsymbol{k}_1}^2(\boldsymbol{Z})\right] + E_{F_{\boldsymbol{Z}}}\left[c_{\boldsymbol{k}_2}^2\Phi_{\boldsymbol{k}_2}^2(\boldsymbol{Z})\right] + \cdots + E_{F_{\boldsymbol{Z}}}\left[c_{\boldsymbol{k}_{\mathbb{Q}}}^2\Phi_{\boldsymbol{k}_{\mathbb{Q}}}^2(\boldsymbol{Z})\right] \\
&\quad + E_{F_{\boldsymbol{Z}}}\left[2c_{\boldsymbol{k}_1}\Phi_{\boldsymbol{k}_1}(\boldsymbol{Z})c_{\boldsymbol{k}_2}\Phi_{\boldsymbol{k}_2}(\boldsymbol{Z})\right] + \cdots \\
&= c_{\boldsymbol{k}_1}^2 + c_{\boldsymbol{k}_2}^2 + \cdots + c_{\boldsymbol{k}_{\mathbb{Q}}}^2 + 0 + \cdots = \sum_{0\le|\boldsymbol{k}|\le \mathrm{K}} c_{\boldsymbol{k}}^2.
\end{aligned} \tag{I6}
$$

Without loss of generality, let $\boldsymbol{k}_1 = \boldsymbol{0}$. For B,

$$
\begin{aligned}
B &= E_{F_{\boldsymbol{Z}}}\left[c_{\boldsymbol{0}}\sum_{|\boldsymbol{k}|\le \mathrm{K}} c_{\boldsymbol{k}}\,\Phi_{\boldsymbol{k}}(\boldsymbol{Z})\right] \\
&= E_{F_{\boldsymbol{Z}}}\left\{c_{\boldsymbol{0}}\Phi_{\boldsymbol{0}}(\boldsymbol{Z})\left[c_{\boldsymbol{0}}\Phi_{\boldsymbol{0}}(\boldsymbol{Z}) + c_{\boldsymbol{k}_2}\Phi_{\boldsymbol{k}_2}(\boldsymbol{Z}) + \cdots + c_{\boldsymbol{k}_{\mathbb{Q}}}\Phi_{\boldsymbol{k}_{\mathbb{Q}}}(\boldsymbol{Z})\right]\right\} \\
&= c_{\boldsymbol{0}}^2 E_{F_{\boldsymbol{Z}}}[\Phi_{\boldsymbol{0}}(\boldsymbol{Z})\Phi_{\boldsymbol{0}}(\boldsymbol{Z})] + c_{\boldsymbol{0}}c_{\boldsymbol{k}_2}E_{F_{\boldsymbol{Z}}}\left[\Phi_{\boldsymbol{k}_2}(\boldsymbol{Z})\Phi_{\boldsymbol{0}}(\boldsymbol{Z})\right] + \cdots + c_{\boldsymbol{0}}c_{\boldsymbol{k}_{\mathbb{Q}}}E_{F_{\boldsymbol{Z}}}\left[\Phi_{\boldsymbol{k}_{\mathbb{Q}}}(\boldsymbol{Z})\Phi_{\boldsymbol{0}}(\boldsymbol{Z})\right] \\
&= c_{\boldsymbol{0}}^2 + 0 + \cdots + 0 = c_{\boldsymbol{0}}^2.
\end{aligned} \tag{I7}
$$

For F,

$$
F = E_{F_{\boldsymbol{Z}}}\left(c_{\boldsymbol{0}}^2\right) = c_{\boldsymbol{0}}^2. \tag{I8}
$$

Substituting A, B, and F into Eq. (I5) yields Eq. (31).

Appendix J. Proof of Proposition 4

Using the least-square method, the discrepancies between the gPCE approximations and simulations are minimized through the following optimization problem:

$$
\min J_e = \sum_{i=1}^{\xi}\left[f\left(\tilde{\theta}_1, \ldots, \tilde{\theta}_i, \ldots, \tilde{\theta}_{N_T+N_P}, \tilde{\phi}_1, \ldots, \tilde{\phi}_i, \ldots, \tilde{\phi}_{N_T+N_P}, \tilde{\zeta}\middle|\boldsymbol{p}^{(i)}\right) - \sum_{j=1}^{\mathbb{Q}} c_{\boldsymbol{k}_j}\,\Phi_{\boldsymbol{k}_j}\left(\boldsymbol{Z}^{(i)}\right)\right]^2. \tag{J1}
$$

Taking the first partial derivative of J_e $w.r.t.$ $c_{\boldsymbol{k}_u}, \forall u \in [1, \mathbb{Q}]$ gives

$$
\frac{\partial J_e}{\partial c_{\boldsymbol{k}_u}} = \sum_{i=1}^{\xi}\left\{\left[f\left(\tilde{\theta}_1, \ldots, \tilde{\theta}_i, \ldots, \tilde{\theta}_{N_T+N_P}, \tilde{\phi}_1, \ldots, \tilde{\phi}_i, \ldots, \tilde{\phi}_{N_T+N_P}, \tilde{\zeta}\middle|\boldsymbol{p}^{(i)}\right) - \sum_{j=1}^{\mathbb{Q}} c_{\boldsymbol{k}_j}\,\Phi_{\boldsymbol{k}_j}\left(\boldsymbol{p}^{(i)}\right)\right]\Phi_{\boldsymbol{k}_u}\left(\boldsymbol{Z}^{(i)}\right)\right\}. \tag{J2}
$$

According to the notations defined in Eqs. (33) to (35), equating the first partial derivatives to zero yields the following matrix form:

$$
\frac{\partial J_e}{\partial c_{\boldsymbol{k}_u}} = \left(\boldsymbol{b}^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{A}\right)\begin{bmatrix}\Phi_{\boldsymbol{k}_u}\left(\boldsymbol{Z}^{(1)}\right)\\ \ldots \\ \Phi_{\boldsymbol{k}_u}\left(\boldsymbol{Z}^{(\xi)}\right)\end{bmatrix} = 0, \forall u \in [1, \mathbb{Q}]. \tag{J3}
$$

Thus,

$$
\left[\frac{\partial J}{\partial c_{\boldsymbol{k}_1}}, \ldots, \frac{\partial J}{\partial c_{\boldsymbol{k}_{\mathbb{Q}}}}\right] = \left(\boldsymbol{b}^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{A}\right)\underbrace{\begin{bmatrix}\Phi_{\boldsymbol{k}_1}\left(\boldsymbol{Z}^{(1)}\right) & \cdots & \Phi_{\boldsymbol{k}_{\mathbb{Q}}}\left(\boldsymbol{Z}^{(1)}\right)\\ \vdots & \ddots & \vdots \\ \Phi_{\boldsymbol{k}_1}\left(\boldsymbol{Z}^{(\xi)}\right) & \cdots & \Phi_{\boldsymbol{k}_{\mathbb{Q}}}\left(\boldsymbol{Z}^{(\xi)}\right)\end{bmatrix}}_{\boldsymbol{A}^{\mathrm{T}}} = \boldsymbol{0}^{\mathrm{T}}. \tag{J4}
$$

Solving Eq. (J4) gives

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}. \tag{J5}$$

Taking the transpose on both sides yields

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{c} = \boldsymbol{A}\boldsymbol{b}. \tag{J6}$$

Left-multiplying by the inverse matrix of $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ yields Eq. (32).

Appendix K. Jacobi polynomials

The Jacobi polynomials in the form of Rodrigues's formula are given by

$$\Phi_k(Z) = P_k^{(a,b)}(z) = \frac{(-1)^k}{2^k k!}(1-z)^{-a}(1+z)^{-b}\frac{d^k[(1-z)^{a+k}(1+z)^{b+k}]}{dz^k}, \forall k \in \mathbb{N}, \tag{K1}$$

where $a, b > -1$, the support is $[-1, 1]$, and the weighting function is

$$w(z) = (1-z)^a(1+z)^b. \tag{K2}$$

The orthogonality of Jacobi polynomials $w.r.t.$ the weighting function $w(z)$ gives

$$\begin{aligned}&\int_{-1}^{1} P_{k_1}^{(a,b)}(z)\, P_{k_2}^{(a,b)}(z) w(z) dz \\ &= \frac{2^{a+b+1}\Gamma(k_1+a+1)\Gamma(k_1+b+1)}{(2k_1+a+b+1)\Gamma(k_1+a+b+1)n!}\delta_{k_1k_2}, \forall k_1, k_2 \in \mathbb{N},\end{aligned} \tag{K3}$$

where $\Gamma(\cdot)$ represents the gamma function defined by

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt; \tag{K4}$$

and $\delta_{k_1k_2}$ is the Kronecker delta function defined by

$$\delta_{k_1k_2} = \begin{cases} 0 & if\ k_1 \neq k_2 \\ 1 & if\ k_1 = k_2 \end{cases}, \forall k_1, k_2 \in \mathbb{N}. \tag{K5}$$

However, the standard beta distribution is defined on the interval $[0, 1]$ with the probability density function

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \tag{K6}$$

where $x \in [0,1]$; $\alpha, \beta > 0$; and $B(\cdot,\cdot)$ represents the beta function defined as

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt. \tag{K7}$$

Therefore, necessary transformations are required to link Jacobi polynomials and the beta distribution. By taking the parameters and variables of the beta distribution as a basis, the parameters and variables in the Jacobi polynomials can be rewritten as

$$\begin{aligned} a &= \beta - 1, \\ b &= \alpha - 1, \\ z &= 2x - 1. \end{aligned} \tag{K8}$$

Substituting Eq. (K8) into Eq. (K3) and letting $k_1 = k_2 = k$ gives

$$\int_0^1 \left[P_k^{(\beta-1,\alpha-1)}(2x-1)\right]^2 w(2x-1)d(2x-1) = \frac{2^{\alpha+\beta-1}\Gamma(k+\alpha)\Gamma(k+\beta)}{(2k+\alpha+\beta-1)\Gamma(k+\alpha+\beta-1)n!}, \forall k \in \mathbb{N}. \tag{K9}$$

Equation (K9) establishes the orthogonality of Jacobi polynomials $w.r.t.$ the weighting function in the transformed parameters and variables. Substituting Eq. (K2) into the left-hand side of Eq. (K9) yields

$$\begin{aligned} &\int_0^1 \left[P_k^{(\beta-1,\alpha-1)}(2x-1)\right]^2 w(2x-1)d(2x-1) \\ &= 2^{\alpha+\beta-1}\int_0^1 \left[P_k^{(\beta-1,\alpha-1)}(2x-1)\right]^2 x^{\alpha-1}(1-x)^{\beta-1}dx \\ &= 2^{\alpha+\beta-1}B(\alpha,\beta)\int_0^1 \left[P_k^{(\beta-1,\alpha-1)}(2x-1)\right]^2 \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}dx \\ &= 2^{\alpha+\beta-1}B(\alpha,\beta)\int_0^1 \left[P_k^{(\beta-1,\alpha-1)}(2x-1)\right]^2 f(x;\alpha,\beta)dx. \end{aligned} \tag{K10}$$

Substituting Eq. (K10) into Eq. (K9) yields

$$\begin{aligned} &\int_0^1 \left[P_k^{(\beta-1,\alpha-1)}(2x-1)\right]^2 f(x;\alpha,\beta)dx \\ &= \frac{\Gamma(k+\alpha)\Gamma(k+\beta)}{(2k+\alpha+\beta-1)\Gamma(k+\alpha+\beta-1)n!\,B(\alpha,\beta)}, \forall k \in \mathbb{N}. \end{aligned} \tag{K11}$$

Equation (K11) indicates that (1) the orthogonal Jacobi polynomials $w.r.t.$ the standard beta distribution, $Beta(\alpha,\beta)$, are $P_k^{(\beta-1,\alpha-1)}(2x-1)$, $\forall k \in \mathbb{N}$; (2) the corresponding normalization constants are $\sqrt{\frac{\Gamma(k+\alpha)\Gamma(k+\beta)}{(2k+\alpha+\beta-1)\Gamma(k+\alpha+\beta-1)n!B(\alpha,\beta)}}, \forall k \in \mathbb{N}$, which ensures the orthogonality of the Jacobi polynomials; and (3) the transformation relationship between the input samples for the original system and the surrogate system is given by $z = 2x - 1$.

Appendix L. Efficiency of Propositions 3 and 4

This section presents the results of a series of numerical experiments comparing the efficiencies of the MCS method and **Propositions 3** and **4** in estimating $E(D)$ and $Var(D)$. As shown in Table L1 and Figure L1, the MCS method required millions of samples to obtain stable estimates for $E(D)$ and $Var(D)$, making it remarkably time-consuming and unsuitable for real-time applications. In contrast, the method based on **Propositions 3** and **4** required only a small number of samples to achieve similar performance. Specifically, when $K = 1$, dozens of samples were sufficient to accurately estimate $E(D)$. Although the estimation of the second-order quantity, $Var(D)$, was more challenging, the proposed method could constrain the APE within 4% using only 210 samples and a $K = 2$ setting. Increasing K to 3 and the number of samples to 1,120 further reduced the error to 0.09%. Moreover, as the proposed method is fully analytical, the computational costs remained at the millisecond level. The excellent

approximation accuracy and high efficiency of the proposed method make it suitable for a wide range of transportation problems.

Table L1. Comparison of the efficiencies of the MCS method and **Propositions 3** and **4**.

Method	Number of samples	$E(D)$ (s)	APE of $E(D)$ (%)	$Var(D)$ (s^2)	APE of $Var(D)$ (%)	Computation time (s)
MCS	10	6164.3	-	651803.3	-	0.00
MCS	100	6472.2	-	507594.5	-	0.01
MCS	1,000	6600.3	-	437981.2	-	0.01
MCS	10,000	6591.5	-	431920.2	-	0.02
MCS	100,000	6585.2	-	443010.9	-	0.28
MCS	1,000,000	6584.2	-	442896.2	-	2.32
MCS (Ground truth)	10,000,000	6584.0	-	442850.2	-	24.06
Propositions 3 and **4** ($K = 1$)	28	6584.1	0.00	314884.6	28.90	0.01
Propositions 3 and **4** ($K = 2$)	210	6586.3	0.04	428672.0	3.20	0.08
Propositions 3 and **4** ($K = 3$)	1,120	6588.5	0.07	443237.8	0.09	0.79

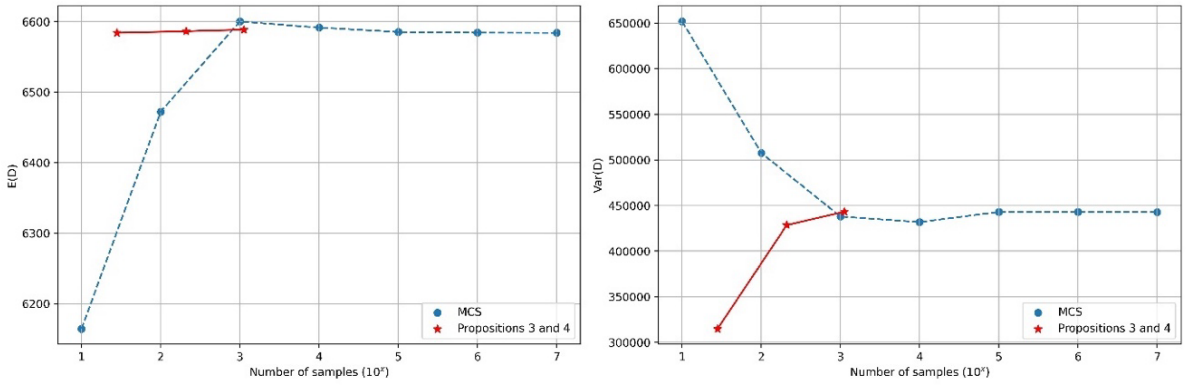


Figure L1. Comparison of the efficiencies of the MCS method and **Propositions 3** and **4**.

Appendix M. Proof of Proposition 5

According to **Proposition 3**,

$$E(D) = \sum_{i=1}^{\xi} a_{1i} f(\theta_1, \dots, \theta_i, \dots, \theta_{N_T+N_P}, \phi_1, \dots, \phi_i, \dots, \phi_{N_T+N_P}, \zeta | \boldsymbol{p}^{(i)}), \tag{M1}$$

$$Var(D) = \sum_{i=2}^{\mathbb{Q}} \left[\sum_{j=1}^{\xi} a_{ij} f\left(\theta_1, \dots, \theta_i, \dots, \theta_{N_T+N_P}, \phi_1, \dots, \phi_i, \dots, \phi_{N_T+N_P}, \zeta | \boldsymbol{p}^{(j)}\right) \right]^2. \tag{M2}$$

For $E(D)$, $\forall x \in \mathbb{V}$,

$$\begin{aligned} \frac{\partial E(D)}{\partial x} &= \lim_{\Delta x \to 0} \frac{\sum_{i=1}^{\xi} \left[a_{1i} f\left(\theta_1, \dots, x + \Delta x, \dots, \zeta | \boldsymbol{p}^{(i)}\right) \right] - \sum_{i=1}^{\xi} \left[a_{1i} f\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(i)}\right) \right]}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{\sum_{i=1}^{\xi} a_{1i} \left[f\left(\theta_1, \dots, x + \Delta x, \dots, \zeta | \boldsymbol{p}^{(i)}\right) - f\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(i)}\right) \right]}{\Delta x} \\ &= \sum_{i=1}^{\xi} a_{1i} \lim_{\Delta x \to 0} \frac{f\left(\theta_1, \dots, x + \Delta x, \dots, \zeta | \boldsymbol{p}^{(i)}\right) - f\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(i)}\right)}{\Delta x} \\ &= \sum_{i=1}^{\xi} a_{1i} f_x'\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(i)}\right) = \mathbb{A}_1 \boldsymbol{b}'. \end{aligned} \tag{M3}$$

For $Var(D)$,

$$\begin{aligned} \frac{\partial Var(D)}{\partial x} = \lim_{\Delta x \to 0} \Bigg\{ & \frac{\sum_{i=2}^{\mathbb{Q}} \left[\sum_{j=1}^{\xi} a_{ij} f\left(\theta_1, \dots, x + \Delta x, \dots, \zeta | \boldsymbol{p}^{(j)}\right) \right]^2 -}{\Delta x} \\ & - \frac{\sum_{i=2}^{\mathbb{Q}} \left[\sum_{j=1}^{\xi} a_{ij} f\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(j)}\right) \right]^2}{\Delta x} \Bigg\}. \end{aligned} \tag{M4}$$

Let $\sum_{j=1}^{\xi} a_{ij} f\left(\theta_1, \dots, x + \Delta x, \dots, \zeta | \boldsymbol{p}^{(j)}\right) = G$ and $\sum_{j=1}^{\xi} a_{ij} f\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(j)}\right) = H$. Eq. (M4) can be rewritten as

$$\frac{\partial Var(D)}{\partial x} = \lim_{\Delta x \to 0} \frac{\sum_{i=2}^{\mathbb{Q}} (G + H)(G - H)}{\Delta x} = \sum_{i=2}^{\mathbb{Q}} \lim_{\Delta x \to 0} \frac{(G + H)(G - H)}{\Delta x}. \tag{M5}$$

For $G + H$,

$$\lim_{\Delta x \to 0} (G + H) = 2 \sum_{j=1}^{\xi} a_{ij} f\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(j)}\right). \tag{M6}$$

According to **Proposition 4**,

$$\sum_{j=1}^{\xi} a_{ij} f\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(j)}\right) = c_{\boldsymbol{k}_i}. \tag{M7}$$

Thus,

$$\lim_{\Delta x \to 0} (G + H) = 2 c_{\boldsymbol{k}_i}. \tag{M8}$$

For $G - H$,

$$\lim_{\Delta x \to 0} \frac{G - H}{\Delta x} = \lim_{\Delta x \to 0} \frac{\sum_{j=1}^{\xi} a_{ij} \left[f\left(\theta_1, \dots, x + \Delta x, \dots, \zeta | \boldsymbol{p}^{(j)}\right) - f\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(j)}\right) \right]}{\Delta x} \tag{M9}$$

$$= \sum_{j=1}^{\xi} a_{ij} \lim_{\Delta x \to 0} \frac{\left[f\left(\theta_1, \dots, x + \Delta x, \dots, \zeta | \boldsymbol{p}^{(j)}\right) - f\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(j)}\right)\right]}{\Delta x}$$

$$= \sum_{j=1}^{\xi} a_{ij} f_x'\left(\theta_1, \dots, x, \dots, \zeta | \boldsymbol{p}^{(j)}\right) = \mathbb{A}_i \boldsymbol{b}'.$$

Substituting Eqs. (M8) and (M9) into Eq. (M5) yields Eq. (39).

Appendix N. Validation of Proposition 5

This appendix provides detailed experimental results of estimations of the gradients of $E(D)$ and $Var(D)$. Numerical differentiation was used to evaluate the gradients of $E(D)$ and $Var(D)$, which were then taken as the ground-truth values for evaluation. To evaluate the partial derivative $\partial E(D)/\partial x, \forall x \in \mathbb{V}$, the expected values at $x + \epsilon$ and $x - \epsilon$, denoted as $E(D')$ and $E(D'')$, respectively, were first computed according to a large number of samples (10^7 samples were adopted), where ϵ was set to 10^{-8}. Then, $\partial E(D)/\partial x$ was estimated as $[E(D') - E(D'')]/(2\epsilon)$. Similar steps were followed to evaluate $\partial Var(D)/\partial x, \forall x \in \mathbb{V}$. The sampling process was conducted only once to save computation time. Additionally, signal groups 10 and 11 (c.f. Figs. 5 and 6) were pedestrian groups and were not included in D; therefore, the partial derivatives $w.r.t.$ the associated variables, θ_{10}, ϕ_{10}, θ_{11}, and ϕ_{11}, were all zeros. A comparison of the results of the numerical differentiation method with those based on **Proposition 5** (Table N1) revealed that the estimates based on **Proposition 5** were nearly identical to those based on numerical differentiation. For the gradient of $E(D)$, the maximum APE was less than 1%, and the average APE was only 0.14%. As expected, the estimation errors for the gradient of $Var(D)$ were slightly larger, with the maximum and average APEs of 4.21% and 2.56%, respectively. Despite the similar performance between the numerical method and **Proposition 5**, **Proposition 5** required a computation time of only 0.08 s, while numerical differentiation required 525.42 s. These results consistently demonstrate the superiority of **Proposition 5**.

Table N1. Comparison of the numerical differentiation method and **Proposition 5** in gradient estimations.

Item	Numerical differentiation (Ground truth)	**Proposition 5**	APE (%)
$\partial E(D)/\partial \theta_1$	5.03	5.04	0.25
$\partial E(D)/\partial \phi_1$	−222.46	−222.46	0.00
$\partial E(D)/\partial \theta_2$	−258.79	−258.81	0.00
$\partial E(D)/\partial \phi_2$	−1045.37	−1045.37	0.00
$\partial E(D)/\partial \theta_3$	139.81	139.75	0.04
$\partial E(D)/\partial \phi_3$	−486.44	−486.44	0.00

$\partial E(D)/\partial\theta_4$	−553.92	−554.18	0.05
$\partial E(D)/\partial\phi_4$	−1060.01	−1060.14	0.01
$\partial E(D)/\partial\theta_5$	1131.42	1130.92	0.04
$\partial E(D)/\partial\phi_5$	−9044.56	−9038.91	0.06
$\partial E(D)/\partial\theta_6$	−1212.09	−1212.17	0.00
$\partial E(D)/\partial\phi_6$	−1923.12	−1923.11	0.00
$\partial E(D)/\partial\theta_7$	90.74	90.71	0.04
$\partial E(D)/\partial\phi_7$	−525.69	−525.69	0.00
$\partial E(D)/\partial\theta_8$	2263.21	2263.07	0.00
$\partial E(D)/\partial\phi_8$	−16057.09	−16055.85	0.00
$\partial E(D)/\partial\theta_9$	−5250.78	−5198.67	0.99
$\partial E(D)/\partial\phi_9$	−12344.84	−12297.99	0.38
$\partial E(D)/\partial\zeta$	−1451272.58	−1438922.88	0.85
Maximum error	-	-	0.99
Average error	-	-	0.14
Variance in error	-	-	0.08
$\partial Var(D)/\partial\theta_1$	−6475.32	−6238.12	3.66
$\partial Var(D)/\partial\phi_1$	−8638.25	−8929.11	3.37
$\partial Var(D)/\partial\theta_2$	−6188.87	−5982.91	3.33
$\partial Var(D)/\partial\phi_2$	−11195.40	−10752.12	3.96
$\partial Var(D)/\partial\theta_3$	3803.32	3774.99	0.74
$\partial Var(D)/\partial\phi_3$	−8725.10	−8465.34	2.98
$\partial Var(D)/\partial\theta_4$	−47673.45	−46060.22	3.38
$\partial Var(D)/\partial\phi_4$	−65062.70	−62852.72	3.40
$\partial Var(D)/\partial\theta_5$	9125.88	9437.14	3.41
$\partial Var(D)/\partial\phi_5$	−96769.76	−100843.69	4.21
$\partial Var(D)/\partial\theta_6$	−23079.59	−23577.60	2.16
$\partial Var(D)/\partial\phi_6$	−30298.61	−31517.06	4.02
$\partial Var(D)/\partial\theta_7$	413.02	419.34	1.53
$\partial Var(D)/\partial\phi_7$	−3254.19	−3270.94	0.51
$\partial Var(D)/\partial\theta_8$	9145.56	9373.16	2.49
$\partial Var(D)/\partial\phi_8$	−88268.94	−90251.22	2.25
$\partial Var(D)/\partial\theta_9$	−777706.30	−771067.34	0.85
$\partial Var(D)/\partial\phi_9$	−875030.54	−893651.21	2.13
$\partial Var(D)/\partial\zeta$	−112443516.26	−112109004.60	0.30
Maximum error	-	-	4.21

Average error	-	-	2.56
Variance in error	-	-	1.51
Computation time (s)	525.42	0.08	-

Appendix O. Traffic demands, successor matrix, and clearance time matrix of Intersection 1

The traffic demands, successor matrix, and clearance time matrix of Intersection 1 in the NGSIM dataset are presented in Tables O1, O2, and O3, respectively.

Table O1. Actual traffic demands and turning proportions of Intersection 1 in the NGSIM dataset.

Direction	Northbound		Southbound		Eastbound		Westbound	
	veh/h	%	veh/h	%	veh/h	%	veh/h	%
Left turn	72	13	80	18	148	22	0	1
Through	480	86	364	81	496	77	68	65
Right turn	0	1	0	1	0	1	36	34
Total	552	1	444	1	644	1	104	1

Table O2. Successor matrix.

Group	1	2	3	4	5	6	7	8	9	10	11
1	-	0	0	0	0	0	-	-	-	-	0
2	1	-	1	1	-	-	1	1	-	1	-
3	1	0	-	1	-	-	1	1	0	-	-
4	1	0	0	-	0	0	-	-	0	-	0
5	1	-	-	1	-	1	1	1	-	1	-
6	1	-	-	1	0	-	1	1	-	-	0
7	-	0	0	-	0	0	-	0	-	-	-
8	-	0	0	-	0	0	1	-	-	-	-
9	-	-	1	1	-	-	-	-	-	1	-
10	-	0	-	-	0	-	-	-	0	-	-
11	1	-	-	1	-	1	-	-	-	-	-

Table O3. Clearance time matrix.

Group	1	2	3	4	5	6	7	8	9	10	11
1	-	6	5	6	6	5	-	-	-	-	6
2	5	-	6	5	-	-	5	6	-	5	-
3	6	5	-	5	-	-	5	5	6	-	-
4	6	6	5	-	5	5	-	-	6	-	5
5	5	-	-	5	-	5	6	5	-	5	-

6	5	-	-	6	6	-	5	5	-	-	6
7	-	5	6	-	5	5	-	6	-	-	-
8	-	5	5	-	5	6	6	-	-	-	-
9	-	-	5	5	-	-	-	-	-	5	-
10	-	5	-	-	5	-	-	-	5	-	-
11	5	-	-	5	-	5	-	-	-	-	-

Appendix P. Sensitivity tests on N_R

This appendix presents the results of preliminary experiments conducted to determine the appropriate parameter N_R. Various cases based on either DPRC or SPRC with different N_R values were considered for signal plan optimizations using the actual demands listed in Table O1 (**Appendix O**). In SPRC, the parameter ω was user-defined to strike a balance between efficiency and stability. According to Eq. (32), $\omega = 0$ signifies that the objective solely optimizes efficiency, while $\omega > 0$ indicates a consideration for both efficiency and stability. A higher ω prioritizes stability and robustness, potentially at the expense of overall efficiency. In addition to cases with $\omega = 0$, $\omega = 20$ was chosen to generate cases that prioritize stability and robustness while maintaining reasonable efficiency. The identical initial signal plan, denoted as $\mathbb{V}^{(0,*)}$, was randomly generated and applied to all of the cases. Delays over 2-h periods across different cases were recorded in VISSIM. The results are summarized in Table P1.

For the DPRC scheme, increasing N_R from 9 to 19 enhanced performance consistently, but further increasing N_R to 29 showed no additional benefit. Thus, $N_R = 19$ was chosen. Similarly, $N_R = 29$ and $N_R = 39$ were determined for the SPRC schemes with $\omega = 0$ and $\omega = 20$, respectively.

Table P1. Sensitivity tests of N_R.

Scheme	ω	N_R	Total junction delay (s)	Chosen?
DPRC	-	9	92,185	
	-	19	91,727	✓
	-	29	92,175	
SPRC	0	9	92,145	
	0	19	91,470	
	0	29	90,041	✓
	0	39	90,398	
	20	9	95,758	
	20	19	93,256	
	20	29	90,755	
	20	39	90,750	✓
	20	49	93,556	

Appendix Q. Properties of SPRC schemes with different ω

This appendix presents additional experiments that compare the properties of the SPRC schemes with ω = 0 and ω = 20. Each experiment was run for a 1,000-min study period, following a 30-min warm-up. To simulate dynamic traffic conditions, the average demand of each approach was modeled using a sinusoidal curve fluctuating by 20% around the constant average demand, starting from an initial phase of zero. That is, while the average demand for each approach varied over time, the mean of these average demands remained identical to the constant average demand used in the warm-up period. The periods of these sinusoidal curves aligned with the duration of the study period. Additionally, the CV penetration rate was set to 0.4. Total junction delays for each cycle over the study period were recorded in VISSIM for analysis. The evaluation metrics were average, maximum, and variance in junction delays. The results are presented in Table Q1.

The results indicate that across different V/C ratios, the SPRC scheme with ω = 20 displayed marginally higher average junction delays but yielded lower maximum junction delays and variances in junction delays than the SPRC scheme with ω = 0. These findings confirm that the SPRC scheme with ω = 0 fully optimized efficiency, whereas the SPRC scheme with ω = 20 provided a balance between optimized efficiency and optimized stability. That is, an increase in ω increased stability at the expense of efficiency.

Table Q1. Comparison of SPRC schemes with different ω.

V/C ratio	SPRC	Average intersection delay (s)	Maximum intersection delay (s)	Variance in intersection delay (s^2)
0.3	$\omega = 20$	561	1,483	27,481
	$\omega = 0$	558	1,496	27,925
0.5	$\omega = 20$	1,111	2,740	102,887
	$\omega = 0$	1,042	3,108	111,105
0.7	$\omega = 20$	5,779	23,987	15,346,807
	$\omega = 0$	5,089	24,232	16,975,453