Available online at www.sciencedirect.com
ScienceDirect
Transportation Research Procedia 7 (2015) 615 – 630
Transportation Research Procedia
www.elsevier.com/locate/procedia

21st International Symposium on Transportation and Traffic Theory

Stochastic Approximations for the Macroscopic Fundamental Diagram of Urban Networks

Jorge A. Laval [a,*], Felipe Castrillón [a]

[a] *School of Civil and Environmental Engineering, Georgia Institute of Technology*

Abstract

This paper proposes a theory for estimating the Macroscopic Fundamental Diagram (MFD) on inhomogeneous corridors and networks using probabilistic methods. By exploiting a symmetry property of the stochastic MFD, whereby it exhibits identical probability distributions in free-flow and congestion, it is found that the network MFD depends mainly on two dimensionless parameters: the mean block length to green ratio and the mean red to green ratio. The theory is validated with an exact traffic simulation and with the empirical data from the city of Yokohama. It is also shown that the effect of buses can be approximated with the proposed theory by accounting for their effect in the red to green ratio parameter.

Selection and peer-review under responsibility of Kobe University
Keywords: Microscopic Fundamental Diagram; Traffic flow; Urban congestion

1. Introduction

It has been shown experimentally by Geroliminis and Daganzo (2008) that the average flow on an urban network can be accurately predicted knowing the average density in the network. This urban-scale Macroscopic Fundamental Diagram (MFD) appears as an invaluable tool to overcome the difficulties of traditional planning models. Although it is still under debate whether it depends on trip origins and destinations and route choice, there is no question that network topology and control parameters such as block length, existence of turn-only lanes, and traffic light settings play a key role.

Existing methods to estimate the MFD analytically for simple homogeneous arterial corridors can be categorized into three types: (i) empirical (Geroliminis and Daganzo, 2007, 2008; Wu et al., 2011; Saberi and Mahmassani, 2012; Geroliminis and Sun, 2011; Geroliminis and Ji, 2011; Cassidy et al., 2011; Knoop, 2012; Gayah and Daganzo, 2011; Buisson and Ladier, 2009; Daganzo et al., 2011), (ii) analytical (Daganzo and Geroliminis, 2008; Leclercq et al., 2014), and (iii) simulation (Ji et al., 2010; Mazlomian et al., 2010; Geroliminis and Boyacı, 2013; Haddad and Geroliminis, 2012; Haddad et al., 2013; Knoop and Hoogendoorn, 2011; Knoop et al., 2011). Existing analytical results are based on the method of cuts for homogeneous corridors, i.e. with equal block size, signal settings and constant offset, and therefore, one can focus on the cuts from a single intersection to compute the MFD for the whole corridor. Despite this apparent simplicity, this approach quickly becomes intractable, more so if buses are

* Corresponding author. Tel. : +1 (404) 894-2360; Fax : +1 (404) 894-2278
E-mail address: jorge.laval@ce.gatech.edu (Jorge A. Laval).

Selection and peer-review under responsibility of Kobe University
doi:10.1016/j.trpro.2015.06.032

introduced (Chiabaut et al., 2014). Even with no buses the homogeneous corridor method cannot be scaled up without complications to estimate the network MFD mainly because a network path cannot be guaranteed to have constant offset all along, even for homogeneous networks.

To overcome these difficulties, in this paper we introduce the concept of stochastic corridors, where any particular inhomogeneous corridor–with different block lengths and signal timing–is seen as a particular realization. Stochastic corridors are in fact probabilistically homogeneous in the sense that the distribution of these network parameters does not change in time or space. This approach allows the estimation of the network MFD, for which analytical methods are currently unavailable.

This paper is organized as follows. Section 2 develops the theory of stochastic corridors, which is based on renewal theory. The existence of short blocks is examined in detail in section 3, as it can severely reduce corridor capacity. Section 4 is devoted to comparing the theory both with an exact traffic simulation and the empirical data from the city of Yokohama presented in Geroliminis and Daganzo (2008). Finally, section 5 presents a discussion.

2. Stochastic corridors

Consider an inhomogeneous corridor consisting of a large sequence of road segments of different length, each one delimited by a traffic signal with settings that vary in time and across segments. This particular corridor is viewed here as a realization of a "stochastic corridor" random variable, where the length of each segment and the red and green times of its signals are random variables ℓ, r and g, respectively, assumed to be independent. We use the symbols μ, σ and $\delta = \sigma/\mu$ for the mean, standard deviation, and coefficient of variation of a random variable, whose name will be indicated as subscript; e.g., block lengths are assumed i.i.d. with mean and variance μ_ℓ, σ_ℓ^2. Turning movements are not considered in our analysis.

We use the superscripts "♯" and "♭" to differentiate variables pertaining to forward and backward cuts, respectively, while the superscript " − " will be used as their placeholder. All links in the network are assumed to obey a triangular fundamental diagram (FD) with free-flow speed $w^\sharp$, wave speed $-w^\flat$ and jam density κ; the saturation flow is therefore $Q = \kappa w^\flat w^\sharp/(w^\flat + w^\sharp)$.

Our formulation is based on variational theory Daganzo (2005a,b), which corresponds to the solution of the kinematic wave model of Lighthill and Whitham (1955); Richards (1956) when expressed as a Hamilton-Jacobi partial differential equation. This solution–known as the Hopf-Lax formula (Lax, 1957; Hopf, 1970)–states that the number of vehicles that have crossed location x by time t, $N(t, x)$, can be expressed in variational form as:

$$N_P = \inf_{B \in \mathcal{B}_P} \{N_B + \Delta_{BP}\} \tag{1}$$

where P is a generic point with coordinates (t, x), $\mathcal{B}_P$ is the set of all points in the boundary that are in the domain of dependence of P, the point $B \equiv (t_B, x_B)$ is in $\mathcal{B}_P$, $N_P \equiv N(t, x)$ and $N_B \equiv N(t_B, x_B)$, and Δ_{BP} is the "cost" or maximum number of vehicles that can cross the minimum path joining B and P; see Fig. 1a. (Notice that in the absence of bottlenecks, such as traffic lights, all valid paths–including the minimum path–between B and P have the same cost and it is customary to compute Δ_{BP} along the straight line BP.)

To derive the corridor MFD, consider the initial value problem in Fig. 1b where the vehicle number $N(t_B, x_B)$ is known in the boundary $t_B = 0$ such that the density, k, is constant. Noting that in this case, $N_B = N_O + (x - x_B)k$ with $N_O \equiv N(0, x)$, we can write

$$N_P - N_O = \min_B \{\Delta_{BP} + (x - x_B)k\}. \tag{2}$$

The MFD is defined as the steady-state flow at any location x; i.e.:

$$q(k) \equiv \lim_{t\to\infty} \frac{1}{t}(N_P - N_O) \tag{3a}$$

$$= \min_B\{\lim_{t\to\infty} \frac{1}{t}(\Delta_{BP} + (x - x_B)k)\}, \tag{3b}$$

$$= \min_v\{\phi(v) + vk\}. \tag{3c}$$

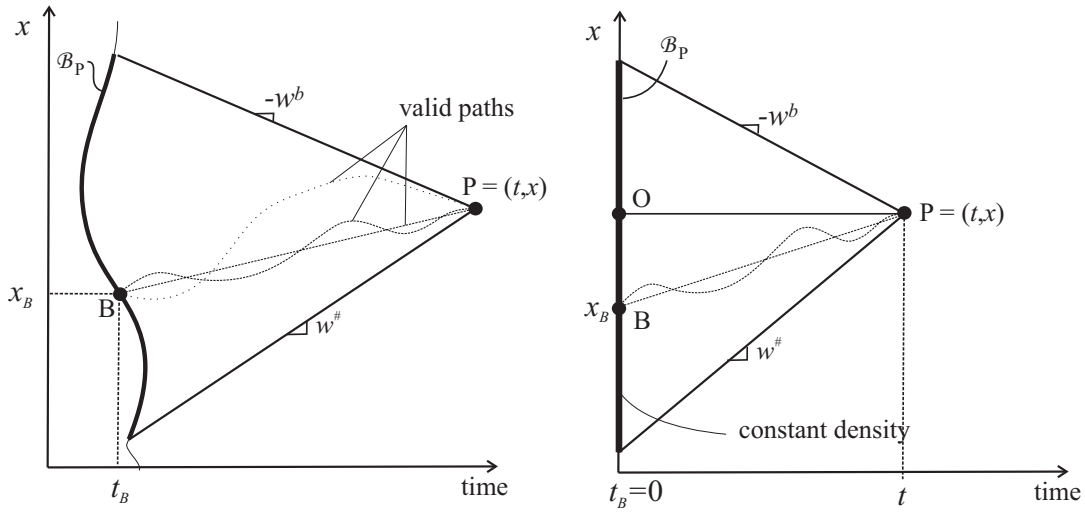

Fig. 1. Definitions of variables in variational theory.

Expression (3c) is the method of cuts (Daganzo and Geroliminis, 2008), where $v \equiv (x - x_B)/t$ is the average speed of the cut and $\phi(v) \equiv \Delta_{BP}/t$ is its maximum passing rate. This method has proven very useful for deterministic homogeneous problems where $\phi(v)$ can be evaluated analytically.

In non-homogeneous corridors the main difficulty is identifying all valid paths between B and P and selecting the one with minimum cost. This problem is exacerbated when the corridor contains short blocks because the minimum path becomes increasingly difficult to calculate. To tackle this problem, this paper proposes using renewal theory to approximate (3b).

When the corridor does not contain short blocks we propose the following approximation. Much as in the method of cuts, the multitude of valid paths in a network is replaced by a small number of "observers" that wander through the network with simple rules with the hope that they will capture the most important valid paths. In our case, we add the restriction that observers must be statistically independent because otherwise the problem becomes intractable.

In the proposed method, observers are emanated at time t and travel backwards in time in a series of "renewal cycles" generated by a strategy, s, until reaching the boundary at some location $x_B = x_B^s$, while having crossed a maximum number of vehicles of $\Delta_{BP} = \Delta_{BP}^s$. In each renewal cycle, the observer starts at the beginning of a red phase, travels for a number of intersections (without crossing a red phase), stops at a given intersection and stays there until the beginning of the previous red phase; see Fig. 2. We define the following strategies, starting at the beginning of a red phase:

- s_0 (stationary observer): stay in the same intersection until reaching the boundary,
- s_1^-: travel at speed w^- backwards in time, stop when hitting a red phase and stay there until the beginning of the red,
- s_2^-: similar to s_1^- but stop at every intersection.

These strategies will generate a total of seven cuts, a number that has been found to be sufficient for the homogeneous case (Leclercq and Geroliminis, 2013). Now, the minimum is taken across s instead of B; setting $x = 0$ without loss of generality, we rewrite (3b) as:

$$q(k) = \min_s \{q_s(k)\}, \qquad \text{where:} \tag{4a}$$

$$q_s(k) = \lim_{t\to\infty} \frac{1}{t} X_B^s(t), \tag{4b}$$

$$X_B^s(t) = \Delta_{BP}^s - x_B^s k, \tag{4c}$$

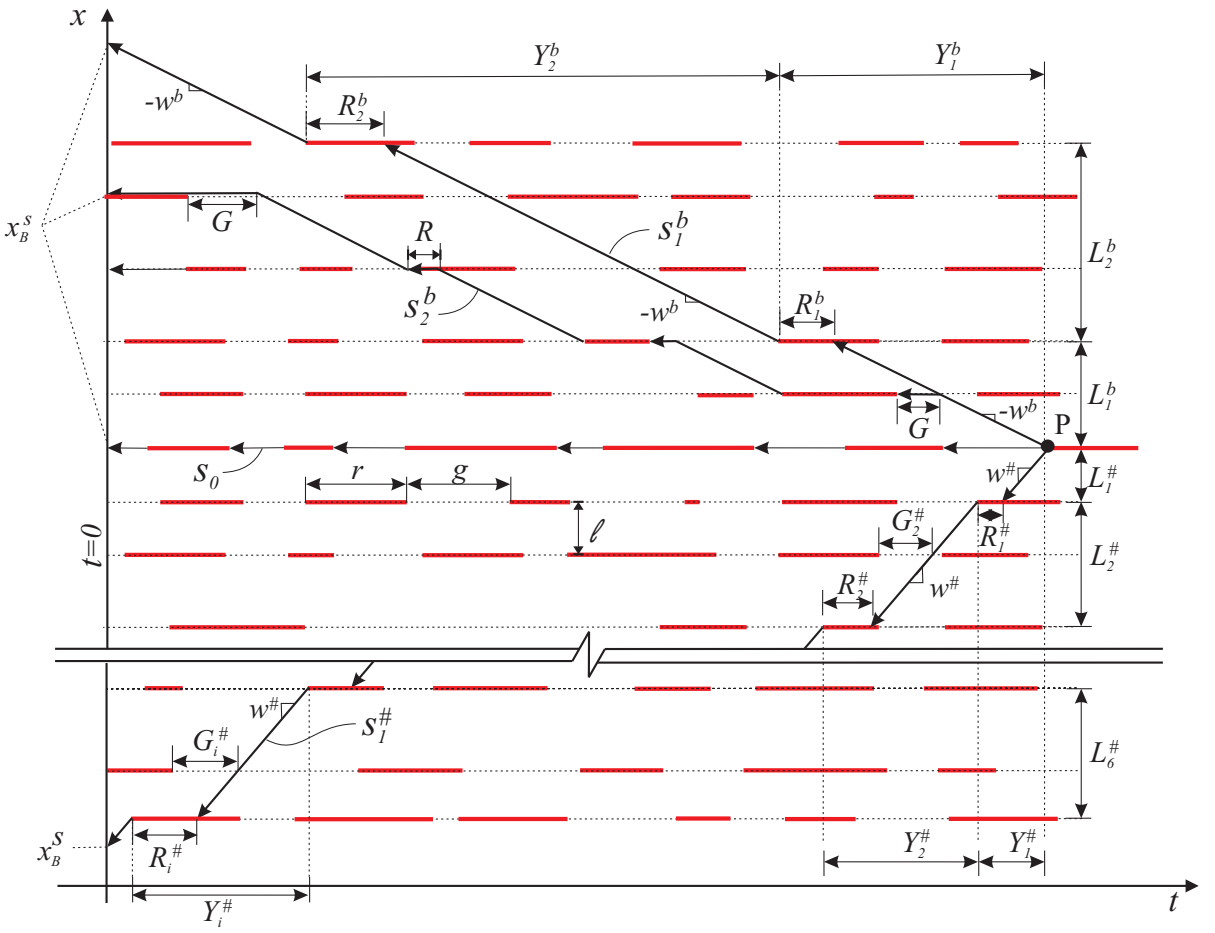

Fig. 2. Definition of strategies in the time space diagram.

where $q_s(k)$ is the cut for strategy s, or the long-term rate of the stochastic process $X_B^s(t)$, which can be written as:

$$X_B^s(t) = \sum_{i=1}^{m(t)} X_i^s, \tag{5}$$

where X_i^s is the contribution (reward) to $X_B^s(t)$ during renewal cycle i and $m^s(t)$ is the number of renewal cycles needed to reach the boundary. As illustrated in Fig. 2, for each renewal cycle i, let Y_i^s be its duration, g_i^s (and r_i^s) the time the observer spends in a green (red) phase, n_i^s the number of intersections crossed, L_i^s the distance traveled. For clarity in notation (i) we have omitted in these definitions the placeholder "–" to differentiate forward and backward cuts, and (ii) hereafter we also omit s-dependencies as much as possible.

The advantage of this formulation is that $X_B(t)$ becomes a renewal reward process since both X_i and Y_i are random variables whose distribution is independent of i; i.e.: $X_i \overset{d}{=} X$ and $Y_i \overset{d}{=} Y$, where the symbol "$\overset{d}{=}$" means "equal in distribution". This is true in our case because, as it will become clear later, we also have that $g_i \overset{d}{=} T$, $r_i \overset{d}{=} r$, $n_i \overset{d}{=} n$, and $L_i \overset{d}{=} L$. The main result of renewal theory is that the limit distribution of q_s in (4b) is Normal and depends on the first moments of (X, Y); i.e.:

Theorem 1. (Distribution of cuts). *The distribution of $q_s(k)$ is normal with mean μ_s and variance σ_s^2/t, where:*

$$\mu_s = \mu_X/\mu_Y \tag{6a}$$

$$\sigma_s^2 = \frac{\mu_X^2}{\mu_Y}(\delta_X^2 + \delta_Y^2 - 2cor(X,Y)\delta_Y\delta_X), \tag{6b}$$

and $cor(X,Y) \equiv Cov(X,Y)/(\sigma_X\sigma_Y)$ is the correlation between X and Y.

Proof. As explained in Serfozo (2009), this result is the application of the Central Limit Theorem to renewal reward processes, which in general states that:

$$\mu_s \equiv E[X_B^s(Y)]/\mu_Y, \tag{7a}$$

$$\sigma_s^2 \equiv V[X_B^s(Y) - \mu_s Y]/\mu_Y, \tag{7b}$$

In our case $X = X_B^s(Y)$ and $V[X - \mu_s Y] = \sigma_X^2 + \mu_s^2\sigma_Y^2 - 2\mu_s \mathrm{Cov}(X,Y)$, and the result follows. □

Note that $\sigma_s^2/\mu_s^2 = \mu_Y((\delta_X - \delta_Y)^2 + 2(1 - \mathrm{cor}(X,Y))\delta_X\delta_Y) > \mu_Y(\delta_X - \delta_Y)^2$ since the correlation $0 < \mathrm{cor}(X,Y) < 1$ in our case. It follows that the coefficient of variation of $q_s(k)$, $\delta_s \equiv \sigma_s/(\mu_s t^{1/2})$, satisfies:

$$\delta_s > |\delta_X - \delta_Y|(\frac{\mu_Y}{t})^{1/2}. \tag{8}$$

This lower bound for the variations in $q_s(k)$ should be tight because $\mathrm{cor}(X,Y) \approx 1$ due to the common terms in X and Y. Also note that δ_s diminishes slowly with time.

To obtain X_i^s, we recall from variational theory with a triangular fundamental diagram that the contribution to Δ_{BP}^s in (4c) is Qg_i for forward cuts and $Qg_i + \kappa L_i$ for backward cuts, while the contribution to the term $x_B^s k$ is $-L_i k$ and $L_i k$, respectively. Therefore,

$$X_i^\sharp = L_i^\sharp k + Qg_i^\sharp, \tag{9a}$$

$$X_i^\flat = (\kappa - k)L_i^\flat + Qg_i^\flat, \qquad \text{while renewal cycle durations are:} \tag{9b}$$

$$Y_i^- = L_i^-/w^- + g_i^- + r_i^-. \tag{9c}$$

It can be seen that the MFD is a function of a large number of parameters: fundamental diagram parameters $w^\sharp, Q, \kappa$ and at least two parameters per random variable ℓ, r and g. To facilitate our analysis we note that the problem can be expressed in terms of only three parameters by reformulating it in dimensionless form and applying a density transformation, as shown next.

2.1. Canonical formulation

In this section we reduce the formulation to the minimum number of parameters. Towards this end, we (i) set all coefficients of variation equal to δ for simplicity, i.e.,

$$\delta \equiv \delta_\ell = \delta_g = \delta_r, \tag{10}$$

(ii) express in dimensionless form the fundamental diagram parameters, (iii) network parameters, and (iv) perform a density transformation.

For the dimensionless formulation of fundamental diagram parameters, let $\theta = w^\sharp/w^\flat$, and measure flow in units of capacity and density in units of jam density. This gives:

$$Q = 1, \qquad \kappa = 1, \qquad w^\sharp = \theta + 1 \qquad \text{and} \qquad w^\flat = (\theta + 1)/\theta. \tag{11}$$

Notice that this rescaling implies, hereafter, that time is measured in units of critical headway and space in units jam spacing. For the network parameters, let ρ be the long-run red to green ratio, and λ be the mean dimensionless block

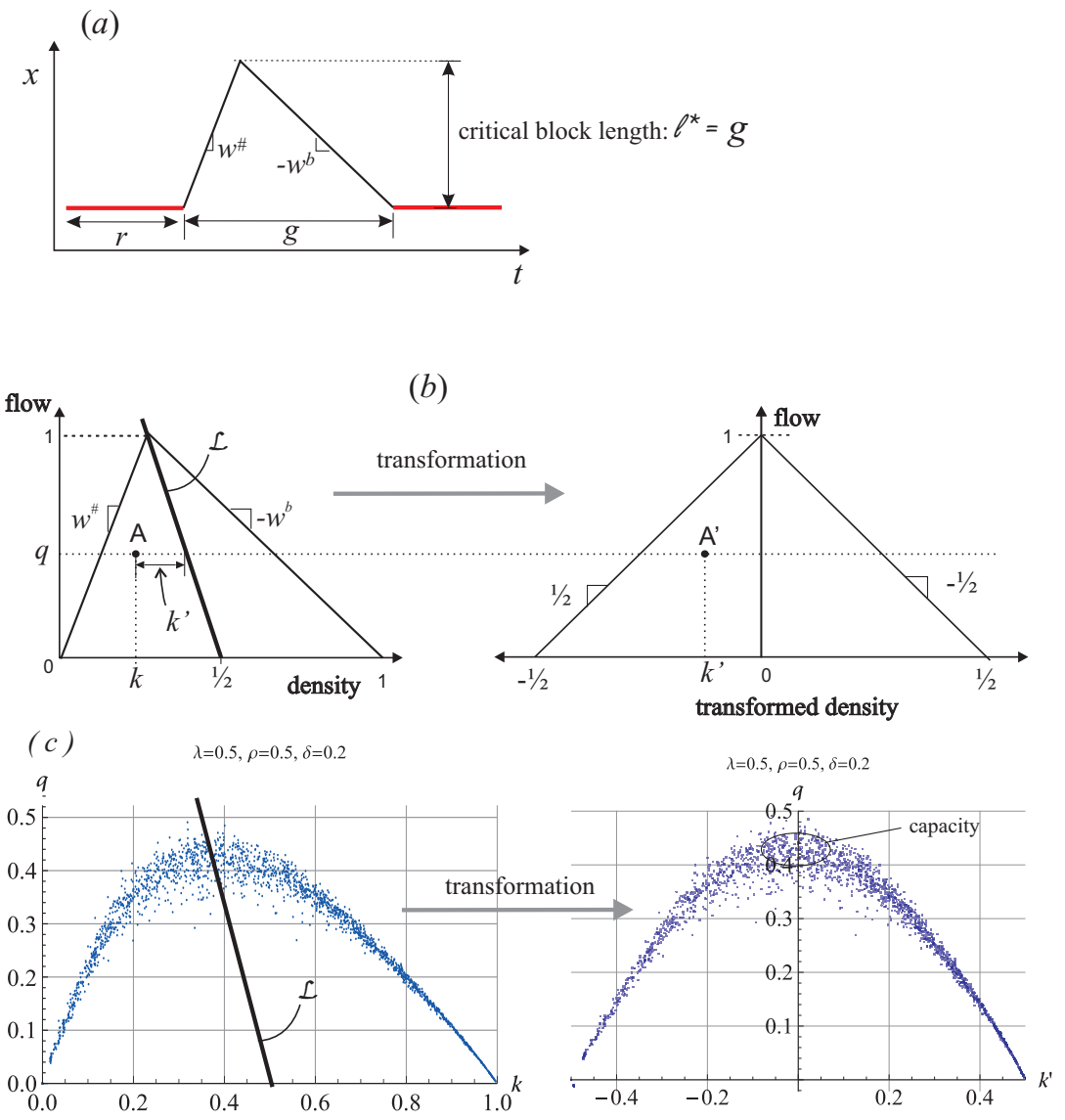

Fig. 3. (a) Critical block length; (b) definition of the density transformation, and (c) an example from simulation.

length measured in units of the mean critical block length $\mu_{\ell^*} \equiv \mu_g/(1/w^\sharp + 1/w^b)$; i.e., $\lambda = \mu_\ell/\mu_{\ell^*}$; see Fig. 3a. As noted in the figure, since $(1/w^\sharp + 1/w^b) = 1$ by (11), we have $\ell^* = g$, and:

$$\rho \equiv \frac{\mu_r}{\mu_g} \quad \text{and} \quad \lambda \equiv \frac{\mu_\ell}{\mu_g}. \tag{12}$$

The density transformation is a mapping $(q, k) \to (q, k')$ that keeps the same flow but expresses density relative to the line $\mathcal{L} : q = (1 - 2k)(1 + \theta)/(\theta - 1)$; see Fig. 3b.

$$k' = k - \frac{1}{2}\left(1 - \left(\frac{1}{w^b} - \frac{1}{w^\sharp}\right)q\right). \qquad \text{(density transformation)} \tag{13}$$

Table 1. Values of the parameters appearing in the RHS of (15) for each strategy. The "=" sign indicates that the entry is equal to the one on the left.

	s_0	s_1	s_2
μ_n	0	$(1+\rho)/\rho$	1
σ_n^2	0	$(1+\rho)/\rho^2$	0
μ_L	0	$\mu_n\mu_\ell$	=
σ_L^2	0	$\mu_n\sigma_\ell^2 + \sigma_n^2\mu_\ell^2$	=
μ_R	μ_r	$\frac{1}{2}\mu_r(1+\delta^2)$	$\frac{1}{2}\left(\rho(1+\delta^2)+2\right)\mu_r/(1+\rho)$
σ_R^2	σ_r^2	$(\delta\mu_R)^2$	$\frac{\left(\delta^2\left(\rho\left(\delta^2+4\rho+14\right)+4\right)+9\rho\right)\mu_r^2}{4(\rho+1)^2}$
μ_G	μ_g	0	$\frac{1}{2}\mu_g(1+\delta^2)/(1+\rho)$
σ_G^2	σ_g^2	0	$\frac{\left(\delta^2\left(\left(\delta^2+6\right)\rho+4\right)+\rho\right)\mu_g^2}{4(\rho+1)^2}$

This transformation has the advantage of making our problems symmetric; see Fig. 3c. Now (9) becomes:

$$X = (1-2|k'|)L/2 + G, \tag{14a}$$

$$Y = L/2 + G + R, \tag{14b}$$

Notice that the transformed density domain is $k' \in [-1/2, 1/2]$. It follows that the first moments of (X, Y) are given by:

$$\mu_X = (1-2|k'|)\mu_L/2 + \mu_G, \qquad \mu_Y = \mu_L/2 + \mu_G + \mu_R, \tag{15a}$$

$$\sigma_X^2 = (1-2|k'|)^2\sigma_L^2/4 + \sigma_G^2, \qquad \sigma_Y^2 = \sigma_L^2/4 + \sigma_G^2 + \sigma_R^2, \tag{15b}$$

$$\mathrm{Cov}(X,Y) = (1-2|k'|)\sigma_L^2/4 + \sigma_G^2. \tag{15c}$$

The values of the parameters appearing in the right-hand side (RHS) of (15) are strategy-specific and are summarized in Table 1. These entries are based on the definition of each strategy; the following observations are in order: (*i*) $n^{s_1} > 0$ is a Geometric random variables with parameter $\rho/(1+\rho)$; (*ii*) $L = \sum_{j=1}^{n} \ell_j$ for s_1 ; (*iii*) the random variables R and G, correspond to g and r for s_0, they correspond to the "age" of a renewal process for s_1, and they are given by a combination of these for $s = s_2$. To see this latter point, notice that when stopping at an intersection, strategy s_1 always lands on a red phase, but strategy s_2 could land on either red or green phases. This explains the strategy-s_2 entries for R and G in the table, as it represents a weighted average of both cases. The derivation for these entries is detailed in Appendix A.

The age of a renewal process is a random variable that represents the time since the last occurrence. In our case, when an observer reaches an intersection during, say, a red phase, it gives the time since the beginning of the red; see Fig. 2. It has known equilibrium distribution. In the case of R this is (G is analogous):

$$F_R(y) \equiv \lim_{t\to\infty} P(R \le y) = \frac{1}{\mu_r}\int_0^y (1 - F_r(z))dz \tag{16}$$

where $F_r(\cdot)$ is the CDF of the red time r. It has also been shown that $\mu_R = \mu_r(1+\delta_r^2)/2$, a result known as the inspection paradox (see e.g. Ross, 2000). The variance, however, depends on F_r; it can be shown that the entries $(\delta\mu_R)^2$ and $(\delta\mu_G)^2$ in Table 1 are upper bounds.

With all, we can now evaluate q_s given in (7) which turns out to be the function of λ, ρ, and δ only. For the mean we have:

$$\mu_{s_0} = \frac{1}{1+\rho}, \tag{17a}$$

$$\mu_{s_1}(k') = \frac{1-2|k'|}{1+c}, \qquad \text{where:} \quad c \equiv (1+\delta^2)\rho^2/(\lambda(1+\rho)), \tag{17b}$$

$$\mu_{s_2}(k') = \frac{1+\delta^2+\lambda(\rho+1)(1\pm 2k')}{\delta^2\left(\rho^2+1\right)+(\rho+1)(\lambda+\rho+1)}. \tag{17c}$$

It can be seen that the mean of s_0 and s_1 depend on a single quantity, ρ and c, respectively, while the mean s_2 depends on all three parameters λ, ρ and δ. It is not difficult to see that the mean flow for all cuts is decreasing in ρ and non-decreasing in λ, as expected. Also, μ_{s_1} (and μ_{s_2}) decreases (increases) approximately linearly with δ^2.

For the squared coefficients of variation, $\delta_s^2 \equiv \sigma_s^2/\mu_s^2$, we get:

$$\delta_{s_0}^2 = \frac{\mu_g}{t}\frac{2\delta^2\rho^2}{\rho+1}, \tag{18a}$$

$$\delta_{s_1}^2 = \frac{\mu_g}{t}\frac{\left(1+\delta^2\right)^2\rho^3\left(1+\delta^2(1+2\rho)\right)}{2(1+\rho)\left((1+\delta^2)\rho^2+(1+\rho)\lambda\right)}, \tag{18b}$$

$$\delta_{s_2}^2(k') = \frac{\mu_g}{t} f(\lambda,\rho,\delta,k'), \tag{18c}$$

where $f(\cdot)$ is a rather intractable rational function given in Appendix B. Examination of (18) both analytically and numerically reveals that δ_s^2 for all strategies is proportional to μ_g/t and increases (approximately) linearly with both δ^2 and ρ; λ-dependencies are none, decreasing, and concave for strategies s_0, s_1 and s_2, respectively. Also note that $\delta_{s_1}^2$ is independent of k', which is unexpected and means that the standard deviation σ_{s_1} is a linear function of k', with maximum at $k' = 0$ and vanishing at $k' = \pm 1/2$.

2.2. Corridor MFD

To compute the corridor MFD we recall that the flow at a given density k' is the minimum of the flow given by each cut. Since the cuts proposed here can be safely assumed to be statistically independent, we have that the CDF of the corridor MFD, $F_{q(k')}(q) = P(q(k') < q)$, can be expressed as:

$$F_{q(k')}(q) = 1 - (1 - F_{s_0}(q))^B \prod_{s\in\Omega}(1 - F_{s,k'}(q)), \tag{19}$$

where $\Omega = \{s_1^\sharp, s_1^\flat, s_2^\sharp, s_2^\flat\}$ and $F_{s,k'}(q)$ is this CDF of the s-cut, given by Theorem 1, i.e. normal distribution with parameters given by (17) and (18). The term $(1 - F_{s_0}(q))^B$ corresponds to the CDF of the minimum of B stationary cuts; this is the only cut requiring this treatment since all other cuts travel at non-zero speeds and therefore visit all intersections in the corridor.

3. Comparison with simulation

In this section we compare the proposed approximations with the traffic simulation described in Appendix C. This simulation is an exact numerical solution of the kinematic wave model of Lighthill and Whitham (1955); Richards (1956) with a triangular fundamental diagram. It can be categorized as a mesoscopic car-following model, whose boundary conditions have been adapted to replicate the problem analyzed in this paper. The corridor is assumed to be a ring road, such that exiting flow immediately reenters the corridor. It is important to note that the simulation requires the same inputs as the theory proposed here: λ, ρ, δ, and also the distribution for ℓ, r and g. Note that the fundamental diagram parameters θ, Q and K needed for the simulation do not affect the results when presented in canonical form.

For each simulation run, we generate (i) the length and (ii) signal settings for each segment as described earlier, and (iii) the initial density to load the corridor at $t = 0$. The output is the average flow over the whole simulation period of T using Edie's definitions Edie (1965). A typical simulation run produces the vehicle trajectories shown on Fig. 4. This process was repeated 8,000 times for each of the 27 combinations of $\lambda = \{0.5, 1, 2\}, \rho = \{0.5, 1, 2\}$ and $\delta = \{0, 0.1, 0.3\}$.

Notice that the case $\delta = 0$ implies deterministic input and therefore a homogeneous corridor, where one would need to define the offset between neighboring traffic signals. Instead, and to make our theory applicable, we set the very first signal cycle start time as uniformly distributed in the time interval $(-(\mu_r + \mu_g), 0)$ for every intersection, which produces a random sequence of offsets.

Hereafter we take advantage of the following theorem to expedite the simulation and analysis by focusing on one side of the transformed fundamental diagram, since the results for the other side are simply a mirror image.

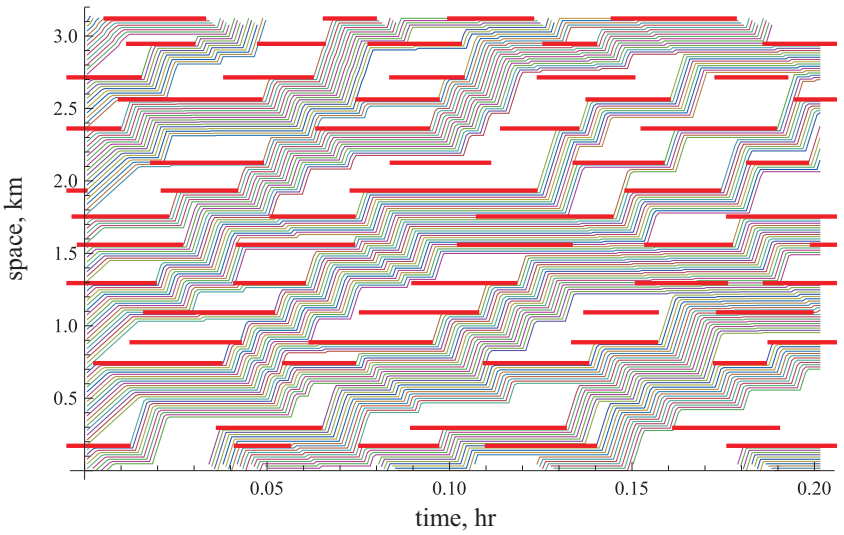

Fig. 4. Example time-space diagram generated by a typical simulation run. Parameter values: $\lambda = 0.5, \rho = 1$ and $\delta = 0.3$.

Theorem 2. (Symmetry). *The distribution of the MFD is symmetric with respect to line $\mathcal{L}$.*

Proof. This is true because under transformation (13): (i) $X_i^- \overset{d}{=} X$ and $Y_i^- \overset{d}{=} Y$, X and Y given in (14), for all renewal cycle i generated under strategy s, and therefore (ii) $\mu_s^\sharp = \mu_s^\flat$ and $\sigma_s^\sharp = \sigma_s^\flat$ in (6). Fact (i) follows from $L_i^\sharp \overset{d}{=} L_i^\flat$, $g_i^\sharp \overset{d}{=} g_i^\flat$, $r_i^\sharp \overset{d}{=} r_i^\flat$, and from the term $(1 - 2|k'|)$ being symmetric with respect to $k' = 0$. □

3.1. Insight from simulation results

The visual inspection of the (q, k) scatter plots produced by the simulation reveals considerable insights; i.e.:

1. only the first two moments of random variables ℓ, r and g have a significant effect in the shape of the MFD, not so the particular probability distribution,
2. the variance in the flow vanishes at $k = 0$ and $k = \kappa$ and appears to be a linearly related to k in the vicinity of these points, which is consistent with the observations regarding (18b),
3. the number of blocks in the corridor, B, has no significant effects in the mean MFD for $B > 5$, only in its variance; see Fig. 5 (top row),
4. the length of the simulation run, T, has no significant effects in the mean MFD for $T > 5$, only slightly in its variance; see Fig. 5 (bottom row),
5. capacity increases with λ and decreases with ρ and δ, as expected, but the effects of λ and ρ are the strongest.

Note that assertions 1 and 2 are based on the visual inspection of the simulation output, which are not included as figures in the paper due to space limitations.

Based on these observations, in the following comparison we use lognormal distributions for ℓ, r and g, $B = 15$ blocks and $T = 15$ min simulation runs.

3.2. Comparison

In 21 of the 27 experiments analyzed here, we can say that model (19) captures the general shape of both the median and the variance the MFD with remarkable accuracy. A sample of these cases is presented in Fig. 6, which shows $q - k'$ diagrams where the dots represent individual simulation runs, and the black curves depict the percentile

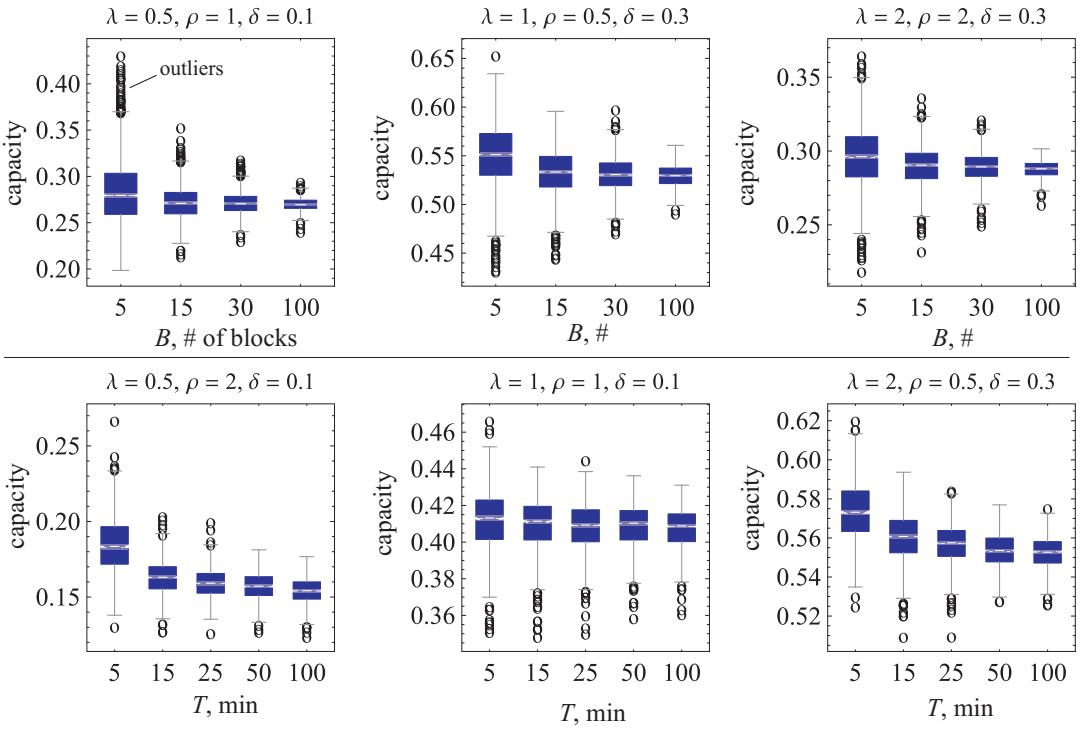

Fig. 5. Box-and-whisker plots for the capacity from the traffic simulation with different number of blocks in the corridor, B. Capacity was estimated only considering data points with a transformed density $k' \in [-0.1, 0.1]$, resulting in a sample size in excess of 1000 in all cases.

curves obtained from (19). Notice that by Theorem 2 we show only one half of the MFD since the other one is a mirror image. The figure also displays the mean s-cuts μ_s from (17); notice that the median predicted by (19) is always below the mean cuts because of the minimum operation involved in (4).

The six cases where discrepancies are observed correspond to very short blocks with moderate red times, i.e. $\lambda = 0.5$ and $\rho \leq 1$. These discrepancies are in the order of 15% in capacity, and can be explained by the effect of short blocks, which is not accounted for explicitly in our model. Interestingly, for long red times the model seems to capture most of this effect as in this case cuts s_1^- and s_2^- remaining well below s_0 near $k' = 0$; see lower-right diagram in Fig. 6.

4. Network MFD ≈ stochastic corridor MFD: evidence from Yokohama

This section describes the notable finding that, in the case of the Yokohama network, if one evaluates the proposed corridor MFD model (19) at the mean values of λ and ρ across the whole network one obtains a surprisingly good approximation of the empirical network MFD reported in Geroliminis and Daganzo (2008). This data consisted of 5-min vehicle counts and occupancy measurements from 500 loop detectors located on most major intersections in a 10 km^2 area in downtown Yokohama, over a weekend day and a weekday in December 2001.

Notice that all parameters needed to evaluate our model were obtained directly from Table 1 in Geroliminis and Daganzo (2008); i.e., the mean values $\lambda = 0.8$ and $\rho = 1.65$, along with the values for θ, Q and K needed to perform the transformation. We used $\delta = 0.2$ but results do not vary significantly with other choices in the range 0.1-0.3.

Fig. 7a presents the original Yokohama data and its transformed version along with the model percentiles in part b of the figure. The good agreement is apparent, especially near capacity. Notice that using this transformation reveals that data points in either sides of the symmetry line overlap nicely in Fig. 7b. This indicates that networks may also subject to the symmetry property put forward in Theorem 2.

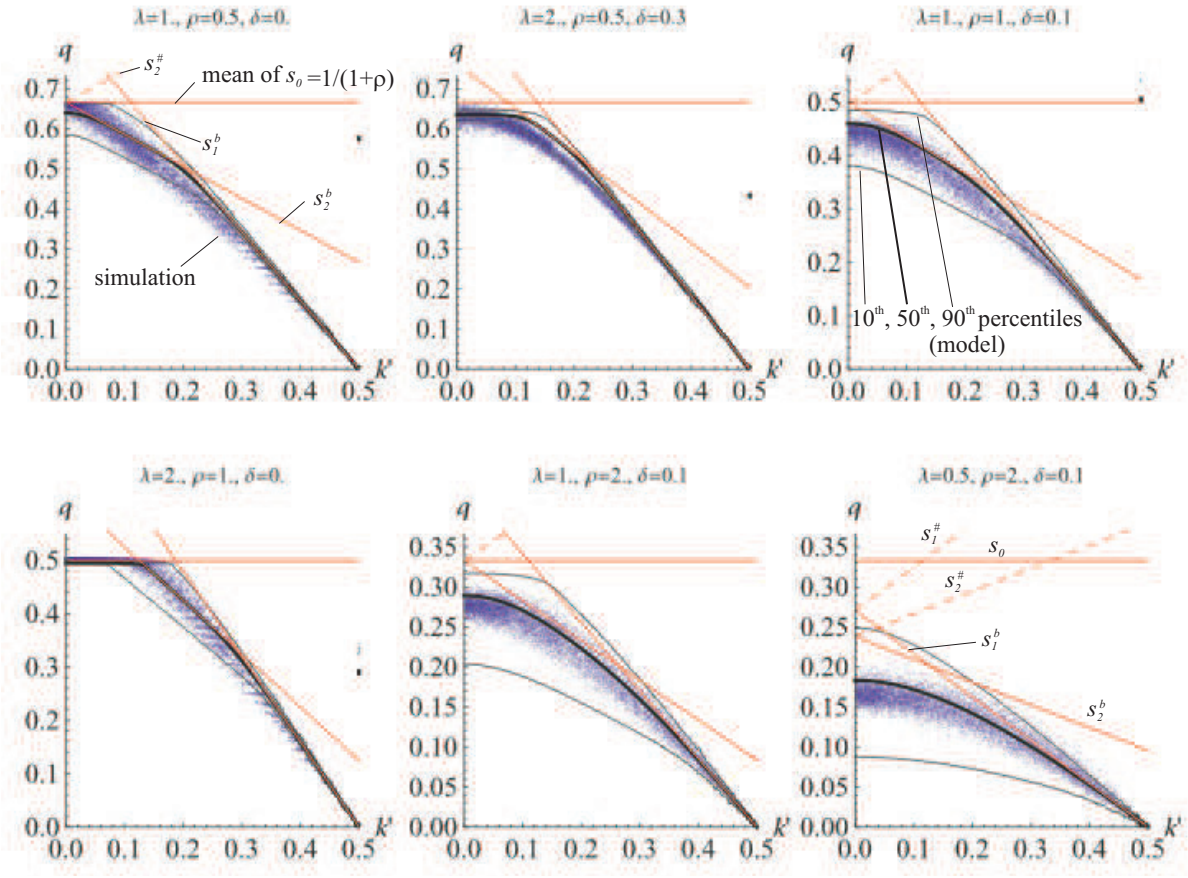

Fig. 6. Comparison between corridor MFD and simulation: $q - k'$ diagrams where dots represent an individual simulation run and the black curves depict the percentile curves obtained from (19). The uniform distribution was used in the simulation.

We conjecture that this good agreement might be explained if, across the Yokohama network, (i) link densities are homogeneous, and (ii) random variables ℓ, g and r are each i.i.d. To see this, consider an alternative definition of the network MFD consisting of flow-density averages across a set of network paths instead of the individual blocks that conform the network. This definition is expected to work well if the density along network paths is approximately constant, and if averages across paths are weighted by this density. These conditions are met when condition (i) is true, and weighted averages can be replaced by simple averages. If in addition (ii) is satisfied, all network paths can be viewed as a realization of a stochastic corridor, and therefore it stands to reason that the stochastic corridor MFD could be a good approximation of the network MFD.

5. Discussion

The main result of this paper is the empirical evidence suggesting that the distribution of the network MFD (i) is a function of parameters λ, ρ and δ defined here, (ii) can be evaluated with the corridor formulas presented here, and (iii) is symmetric when using transformation (13). This result is important because it means that real urban networks can be well approximated by estimating only three observable parameters. This opens the door for a simplified way to analyze complex urban networks, and enables the application of existing control algorithms that rely on the knowledge of the network MFD.

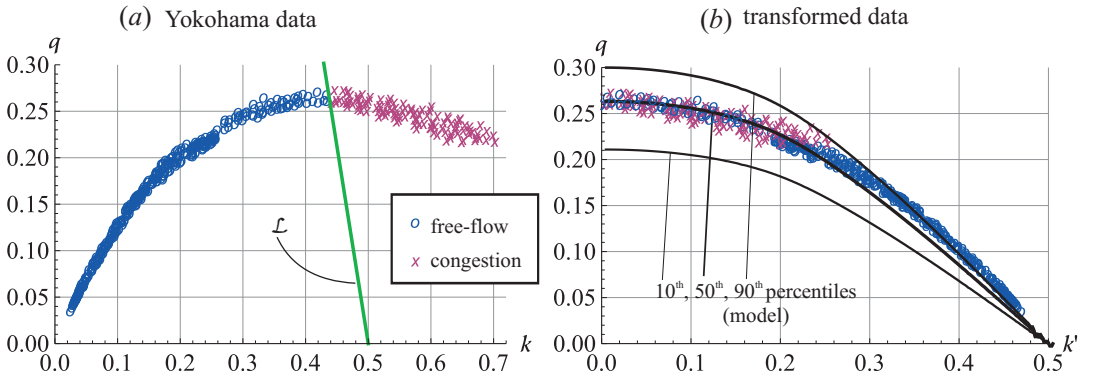

Fig. 7. Yokohama data and comparison with simulation results. Table 1 in Geroliminis and Daganzo (2008) was used to derive $\lambda = 0.8$ and $\rho = 1.65$, along with the values for θ, Q and K needed to perform the transformation. We use $\delta = 0.2$.

We have seen that corridor capacity increases with λ and decreases with ρ and δ, as expected. This suggests that a practical solution to gridlock due to short blocks is reducing cycle times, as this would increase λ while keeping ρ approximately constant. Of course, one should also consider the effect of lost times, which are absent in our framework. We can also say that designing cities with square blocks of constant length, and-or using a constant split ratio on main arterials should be beneficial since in these cases the value of δ would decrease.

The following paragraphs address some of the limitations of the proposed framework, which are being investigated by the authors.

The cuts proposed here were chosen to minimize the overlap of the corresponding observers in the (t, x) plane, to avoid statistical dependencies due to path overlaps. Additional cuts can be defined with the framework proposed here but chances are that they will be correlated. In such case, instead of the MFD distribution formula (19) one should use its general counterpart that involves the joint distribution of all cuts, which may be challenging to identify. Notice, however, that even for the few cuts proposed here independence cannot be guaranteed for all parameter values; e.g., when ρ is very large, observers s_1 and s_2 would be identical since both would stop at every intersection.

For simplicity in the model analysis, we assumed starting in section 2.1 that all coefficients of variations are equal. While this assumption may be violated in practice, it is expected that it will still provide a good approximation because we have seen that λ and ρ are the parameters that affect the shape of the MFD the most. However, from an analytical perspective this assumption is not necessary and closed-form expressions can also be obtained, albeit more complicated.

The existence of turning movements is not considered in the present model. However, the good agreement with the Yokohama MFD suggests that this is a reasonable approximation. The case of heavy turning movements could be tackled with the Hamilton-Jacobi formulations including source terms.

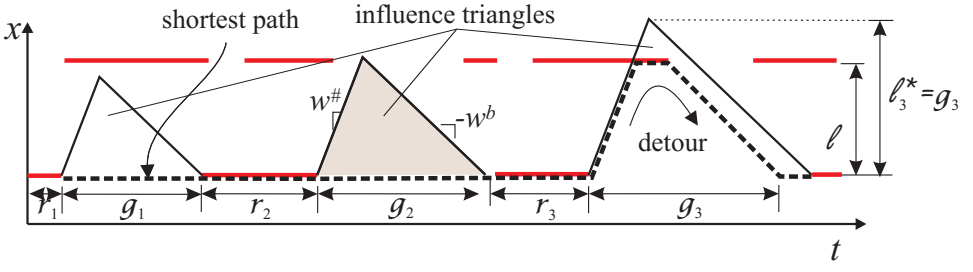

Fig. 8. Detours a corridor with short blocks.

The effect of short blocks remains a challenge. They can alter the minimum paths in ways not captured by the model proposed here. In particular, the stationary cut would no longer be a straight-line observer, but one that "takes detours" through several neighboring intersections and eventually comes back; see Fig. 8a. A detour will be taken by the minimum path each time a red phase on the neighboring intersection falls inside the "influence triangle" of the green phase on the original intersection; see Fig. 8. The height of this triangle is the critical block length $\ell^* = g$, and therefore a detour can happen if $\ell < g$. Since the random variable $\ell^* - g$ has mean $\mu_g(\lambda - 1)$ and variance $\mu_g^2\delta^2(\lambda^2 + 1)$, it can be shown that, assuming normality and to a first-order approximation, if

$$\lambda < 1 + \delta, \qquad \text{(short block condition)}, \tag{20}$$

the probability of detours is non-negligible. The magnitude of the corresponding capacity loss would also depend on ρ, and we saw in section 3.2 that our model is able to capture the capacity losses even when (20) is violated, except in the cases $\lambda = 0.5$ and $\rho \leq 1$, which remain elusive. In the meantime, the following regression model gives the mean MFD capacity of all the experiments simulated in this paper:

$$\text{mean capacity} = \frac{1}{1 + \rho\,(0.58\delta\lambda + 1.64\lambda^2 - 5.3\lambda + 4.99)}, \tag{21}$$

with an adjusted $R^2 = 0.997$ and a standard error of 0.026. Notice that this expression should give the universal shape of network capacity and does not require recalibration. This is true because all parameters are dimensionless and they are the only ones appearing in the formulation proposed in this paper.

In closing, we conjecture that the effect of buses can be approximated with the proposed theory by accounting for their effect in ρ, even for homogeneous networks. To illustrate this, Fig. 9 presents simulation results of a homogeneous network with constant block length and signal settings, $\lambda = 1, \rho = 1$ and zero offset, (a) without buses and (b) with buses running at 2min headways in 10% of the streets making regular stops just upstream of every intersection. As expected, the MFD without buses is not symmetric and exhibits no variance. However, the presence of buses has a smoothing effect restoring symmetry to the MFD. This can be explained by the link between stationary and moving bottlenecks (Laval and Daganzo, 2006; Chiabaut et al., 2014), which means that buses have the same effect of a sequence of traffic lights, and therefore they change the distribution of the red to green ratio in the network. We are currently working on an analytical method to obtain the "adjusted" ρ value given the operational parameters of the bus fleet. The preliminary result in part b of the figure is encouraging, where it can be seen that the model proposed here fits the simulation data well with the "adjusted" ρ value.

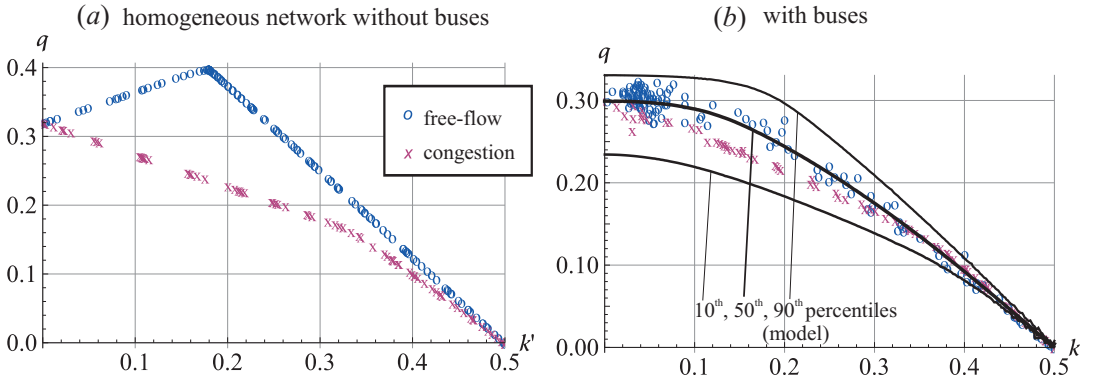

Fig. 9. The effect of buses. Simulation results of a homogeneous network with $\lambda = 1, \rho = 1$ and zero offset, (a) without buses and (b) with buses running at 2min headways in 10% of the streets making regular stops just upstream of every intersection. In part b the model percentiles from (19) assume $\lambda = 1$, "adjusted" $\rho = 1.7$ and $\delta = 0.2$.

Acknowledgments.

This research was supported by NSF research project # 1301057. The authors are grateful to Nikolas Gerolominis, Ludovic Leclercq and Nicolas Chiabaut for stimulating discussion, to Eric Gonzales for providing the Yokohama data, to Yi Zhou for her help with the artwork, and to two anonymous referees, whose comments and suggestions greatly improved the quality of the paper.

Appendix A. Derivation of R and G for strategy s_2

As mentioned in the main text, when stopping at an intersection, strategy s_3 can land on either a red or green phase. Here we are interested in the first moments of the time the observer spends on a red and green phase during a renewal cycle, R and G, respectively. Let R^* and G^* be the time between the landing point and the beginning of the corresponding phase, i.e. the age of a renewal process with distribution given by (16). Let $p = \rho/(1+\rho)$ be the probability that the observer lands on a red, and let I be the indicator variable that equals 1 when the observer lands on a red, and 0 if it lands on a green. Therefore, $E[I] = p$ and $V[I] = p(1-p)$. We have:

$$R = R^* I + r(1-I), \tag{A.1a}$$
$$G = G^*(1-I) + 0I, \tag{A.1b}$$

with means given by

$$\mu_R = \mu_{R^*} p + \mu_r(1-p), \tag{A.2a}$$
$$\mu_G = \mu_{G^*}(1-p), \tag{A.2b}$$

and variances:

$$\sigma_R^2 = \sigma_{R^*}^2 p^2 + (1-p)^2\sigma_r^2 + p(1-p)(E[(R^*)^2] + E[r^2] + 2\mu_r\mu_{R^*}), \tag{A.3a}$$
$$\sigma_G^2 = \sigma_{G^*}^2(1-p)^2 + E[(G^*)^2]p(1-p), \tag{A.3b}$$

Where we have used the conditional variance formula, e.g. $\sigma_G^2 = E[V[G|G^*]] + V[E[G|G^*]]$. In addition, for (A.3a) one needs $\mathrm{Cov}(IR^*, r(1-I) = \mu_r\mu_{R^*}V[I]$. Replacing the terms in the RHS of (A.2) and (A.3) by the expressions given in the main text results in the entries for R and G in Table 1.

Appendix B. Function $f(\lambda, \rho, \delta, k')$

For completeness, here we give expression for $f(\lambda, \rho, \delta, k')$ in (18c):

$$f(\lambda,\rho,\delta,k') = \frac{\rho^2 A + 4k'^2\lambda^2(\rho+1)^2 B + 4|k'|\lambda(\rho+1)\rho C}{2(\rho+1)(\delta^2(\rho^2+1) + (\rho+1)(\lambda+\rho+1))(\delta^2 + \lambda(\rho+1)(1-2|k'|)^2} \tag{B.1}$$

where:

$$\begin{aligned} A =& \delta^{10}\rho(\rho+1)^2 + \delta^8\rho(2\lambda(\rho+1)^2 + 5\rho^2 + 12\rho + 9) + \delta^6\cdot \\ & (2\lambda^2\rho^4 + (5\lambda^2 + 6\lambda + 10)\rho^3 + 2(2\lambda^2 + 7\lambda + 14)\rho^2 + (\lambda^2 + 8\lambda + 34)\rho + 8) + \delta^4\cdot \\ & (\lambda^2\rho(4\rho+7)(\rho+1)^2 + \lambda(6\rho^3 + 34\rho^2 + 36\rho + 8) + 2(5\rho^3 + 16\rho^2 + 29\rho + 8)) + \delta^2\cdot \\ & (\lambda^2(\rho+1)^2(2\rho^2 + 15\rho + 8) + 2\lambda(\rho^3 + 21\rho^2 + 24\rho + 4) + 5\rho^3 + 18\rho^2 + 45\rho + 8) + \rho\cdot \\ & (9\lambda^2(\rho+1)^2 + 18\lambda(\rho+1) + \rho^2 + 4\rho + 13) \end{aligned}$$

$$\begin{aligned} B =& \delta^6(2\rho^4 + \rho^3 + 2\rho^2 + \rho + 2) + \delta^4(4\rho^4 + 7\rho^3 + 4\rho^2 + 7\rho + 4) + \delta^2\cdot \\ & (2\rho^4 + 15\rho^3 + 10\rho^2 + 7\rho + 2) + 9\rho^3 + \rho \end{aligned}$$

$$
\begin{aligned}
C =& \delta^8\rho(\rho^2-1) + \delta^6(\rho+1)(2\lambda\rho^3 + (\lambda+3)\rho^2 + (\lambda-3)\rho - 2) + \delta^4\cdot \\
& (\lambda(4\rho^4 + 11\rho^3 + 9\rho^2 + 4\rho + 2) + 3\rho^3 + 8\rho^2 - 5\rho - 4) + \delta^2\cdot \\
& (\lambda(2\rho^4 + 17\rho^3 + 24\rho^2 + 11\rho + 2) + \rho^3 + 16\rho^2 - 3\rho - 2) + \rho(9\lambda\rho^2 + (9\lambda+8)\rho - 2)
\end{aligned}
$$

Appendix C. Traffic simulation

This traffic simulation used in this paper is mesoscopic car-following model that gives the exact numerical solution of the kinematic wave model of Lighthill and Whitham (1955); Richards (1956) with a triangular fundamental diagram, in terms of the position of vehicle n at time t, $X(t,n)$. Here we use the lattice implementation as described in Laval and Leclercq (2013), where vehicle number and time are discretized in increments of Δn and Δt. A "tilde" denotes a dimensionless quantity, e.g. $\widetilde{t} = t/\Delta t$. This gives the lattice X-model in Laval and Leclercq (2013):

$$
X_{\widetilde{t},\widetilde{n}} = \min\left\{X_{\widetilde{t}-1,\widetilde{n}} + \theta\Delta x, X_{\widetilde{t}-1,\widetilde{n}-1} - \Delta x\right\} \tag{C.1}
$$

To keep the solution exact, θ has to be an integer and $\Delta n = \kappa\Delta x, \Delta t = \Delta x/w$. The increment Δx is arbitrary and gives flexibility to this model; it is set such that smallest segments in our simulation has at least eight spatial increments, i.e. $\Delta x = \max[1/\kappa, \ell_{min}/8]$. The parameters used on the simulation are: $w = 20$ km/hr, $\kappa = 150$ veh/km, $\theta = 4$, $\mu_\ell = .2$ km, number of signals $B = 15$, and simulation time $T = 15$ min. The remaing variables are therefore: $\mu_g = \mu_\ell(1/w^\sharp + 1/w^\flat)/\lambda$ and $\mu_r = \mu_g\rho$. Finally, we calculate the average flow over the whole simulation period using Edie's definitions Edie (1965); i.e.:

$$
q = \frac{\Delta n}{LTQ}\sum_{\widetilde{n}}(X_{\widetilde{T},\widetilde{n}} - X_{0,\widetilde{n}}). \tag{C.2}
$$

References

Buisson, C., Ladier, C., 2009. Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundaental diagrams. Transportation Research Record 2124, 127–136.

Cassidy, M., Jang, K., Daganzo, C., 2011. Macroscopic fundamental diagram for freeway networks: Theory and observation, 90th Annual Meeting of the Transportation Research Board, Washington D.C.

Chiabaut, N., Xiea, X., Leclercq, L., 2014. Performance analysis for different designs of a multimodal urban arterial. Transportmetrica B: Transport Dynamics 2, 229–245. doi:DOI: 10.1080/21680566.2014.939245.

Daganzo, C.F., 2005a. A variational formulation of kinematic wave theory: basic theory and complex boundary conditions. Transportation Research Part B 39, 187–196.

Daganzo, C.F., 2005b. A variational formulation of kinematic waves: Solution methods. Transportation Research Part B: Methodological 39, 934 – 950. doi:DOI: 10.1016/j.trb.2004.05.003.

Daganzo, C.F., Gayah, V.V., Gonzales, E.J., 2011. Macroscopic relations of urban traffic variables: Bifurcations, multi-valuedness and instability. Transportation Research Part B Methodological 45, 278–288.

Daganzo, C.F., Geroliminis, N., 2008. An analytical approximation for the macroscopic fundamental diagram of urban traffic. Transportation Research Part B: Methodological 42, 771–781.

Edie, L.C., 1965. Discussion of traffic stream measurements and definitions, in: 2nd Int. Symp. on Transportation and Traffic Theory, Paris, France. p. 139154.

Gayah, V., Daganzo, C., 2011. Clockwise hysteresis loops in the macroscopic fundamental diagram: An effect of network instability. Transportation Research Part B .

Geroliminis, N., Boyacı, B., 2013. The effect of variability of urban systems characteristics in the network capacity. Transportation Research Part B 46(10), 1576–1590.

Geroliminis, N., Daganzo, C., 2007. Macroscopic modeling of traffic in cities, 86th Annual Meeting of the Transportation Research Board, Washington D.C.

Geroliminis, N., Daganzo, C., 2008. Existence of urban-scale macroscopic fundamental diagrams: some experimental findings. Transportation Research Part B: Methodological 42, 759–770.

Geroliminis, N., Ji, Y., 2011. Spatial and temporal analysis of congestion in urban transportation networks, 90th Annual Meeting of the Transportation Research Board, Washington D.C.

Geroliminis, N., Sun, J., 2011. Properties of a well-defined macroscopic fundamental diagram for urban traffic. Transportation Research Part B 45, 605–617.

Haddad, J., Geroliminis, N., 2012. On the stability of traffic perimeter control in two-region urban cities. Transportation Research Part B 46(9), 1159–1176.

Haddad, J., Ramezani, M., Geroliminis, N., 2013. Cooperative traffic control of a mixed network with two urban regions and a freeway. Transportation Research Part B 54, 17–36.

Hopf, E., 1970. On the right weak solution of the cauchy problem for a quasilinear equation of first order. Indiana Univ. Math. J. 19, 483–487.

Ji, Y., Daamen, W., Hoogendoorn, S., Hoogendoom-Lanser, S., Qian, X., 2010. Investigating the shape of the macroscopic fundamental diagram using simulation data. Transportation Research Record: Journal of the Transportation Research Board , 40–48.

Knoop, V., Hoogendoorn, S., 2011. Variable macroscopic fundamental diagrams for traffic networks, 9th Conference on Traffic and Granular Flow, Springer (in print).

Knoop, V., Lint, J.V., Hoogendoorn, S., 2011. Route advice and its effect on the macroscopic fundamental diagram, 9th Conference on Traffic and Granular Flow, Springer (in print).

Knoop, V.L., 2012. The impact of traffic dynamics on the macroscopic fundamental diagram, 92th Annual Meeting of the Transportation Research Board, Washington D.C.

Laval, J.A., Daganzo, C.F., 2006. Lane-changing in traffic streams. Transportation Research Part B 40, 251–264.

Laval, J.A., Leclercq, L., 2013. The Hamilton-Jacobi partial differential equation and the three representations of traffic flow. Transportation Research Part B 52, 17–30. doi:10.1016/j.trb.2013.02.008.

Lax, P.D., 1957. Hyperbolic systems of conservation laws ii, in: Sarnak, P., Majda, A. (Eds.), Communications on Pure and Applied Mathematics. Wiley Periodicals, p. 537566. doi:10.1002/cpa.3160100406.

Leclercq, L., Chiabaut, N., Trinquier, B., 2014. Macroscopic fundamental diagrams: A cross-comparison of estimation methods. Transportation Research Part B: Methodological 62, 1 – 12. doi:http://dx.doi.org/10.1016/j.trb.2014.01.007.

Leclercq, L., Geroliminis, N., 2013. Estimating {MFDs} in simple networks with route choice. Transportation Research Part B: Methodological 57, 468 – 484. doi:http://dx.doi.org/10.1016/j.trb.2013.05.005.

Lighthill, M.J., Whitham, G., 1955. On kinematic waves. I Flow movement in long rivers. II A theory of traffic flow on long crowded roads. Proceedings of the Royal Society of London 229, 281–345.

Mazlomian, A., Geroliminis, N., Helbing, D., 2010. The spatial variability of vehicle densities as determinant of urban network capacity. Philosophical Transactions of the Royal Society A 368, 4627–4647.

Richards, P.I., 1956. Shockwaves on the highway. Operations Research , 42–51.

Ross, S., 2000. Introduction to Probability Models. 7th ed., Academic Press.

Saberi, M., Mahmassani, H.S., 2012. Exploring the properties of network-wide flow-density relations in a freeway network. Transportation Research Record , 153–163.

Serfozo, R., 2009. Basics of Applied Stochastic Processes. Springer-Verlag Berlin Heidelberg.

Wu, X., Liu, H., Geroliminis, N., 2011. An empirical analysis on the arterial fundamental diagram. Transportation Research Part B: Methodological 45(1), 255–266.