IOP Publishing Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **46** (2013) 445101 (19pp) doi:10.1088/1751-8113/46/44/445101

Theoretical predictions for vehicular headways and their clusters

Milan Krbálek

Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Prague, Czech Republic

E-mail: milan.krbalek@fjfi.cvut.cz

Received 17 August 2013
Published 18 October 2013
Online at stacks.iop.org/JPhysA/46/445101

Abstract

This paper presents a derivation of analytical predictions for steady-state distributions of netto time gaps among clusters of vehicles moving inside a traffic stream. Using the thermodynamic socio-physical traffic model with short-ranged repulsion between particles (originally introduced in Krbálek and Helbing 2004 *Physica* A **333** 370) we first derive the time-clearance distribution in the model and confront it with relation to the theoretical criteria for the acceptability of analytical clearance distributions. Consecutively, the approximating statistical distributions for the so-called time multi-clearances are calculated by means of the theory of functional convolutions. Moreover, all the theoretical surmises used during the above-mentioned calculations are evaluated by the statistical analysis of traffic data. The mathematical predictions acquired in this paper are thoroughly compared with relevant empirical quantities and discussed in the context of traffic theory.

PACS numbers: 05.40.−a, 89.40.−a, 05.45.−a

(Some figures may appear in colour only in the online journal)

1. Introduction

Explorations of traffic micro-quantities (i.e. quantities belonging to individual vehicles) and efforts to predict their statistics are as old as traffic research itself. Virtually, all important reviews on traffic science (for example [1–3], or [4]) try to explain at least some basic knowledge on statistical distributions of time intervals or distance-gaps among moving cars. Although the theoretical prediction of evolution for headway distributions is still extremely vague, some partial achievements have been reached in the last few years (for example [5–8, 11, 12, 14–18], or [19]). We try to pick up the threads of those results and get closer to the heart of the matter.

This contribution is focused predominantly on the evolution of netto time intervals (so-called *time clearances*) between succeeding cars passing a given point (a traffic detector, typically) located at an expressway. Investigations of the traffic clearances are significantly

advantageous (contrary to explorations of distance-gaps) because of their direct measurability. Indeed, most traffic detectors gauge the time of a vehicle's passage directly, which means that those data do not show any systematic or mediated errors (provided the detector is not damaged). Except time clearances, we concentrate our attention on so-called *multi-clearances*, i.e. cumulated time clearances among several consecutive cars. By analogy, such a quantity is (similarly to the time clearance) directly measurable, which opens a possibility for detailed statistical investigations of multi-clearance distributions with respect to the location of the traffic ensemble in the fundamental diagram.

In fact, the empirical and/or theoretical investigations of traffic multi-headways are not sporadic in the physics of traffic. Some of the previous scientific works, e.g. [23–26], or [27], briefly analyze the multi-clearance distributions (predominantly from the empirical point of view) or their statistical variances. The reasons for investigations of such a type are obvious, since a deeper comprehension of changes in traffic microstructure will provide a more thorough insight into convoluted traffic interactions.

2. Preliminary evaluation of empirical observations

The vehicular-data records analyzed in this paper are typically of the following types. The set

$$T_\ell^{(\mathrm{in})} = \left\{\tau_{k\ell}^{(\mathrm{in})} \in \mathbb{R}^+ | k = 1, 2, \ldots, N\right\}, \qquad (\ell \in \{0, 1, 2\}) \tag{1}$$

includes the chronologically-ordered times when the front bumper of kth car has intersected the detector line (located at ℓth lane of an expressway). The respective times $\tau_{k\ell}^{(\mathrm{out})}$ when the rear bumper of kth car has intersected the detector line are summarized in the set $T_\ell^{(\mathrm{out})}$. Analogously, the set of velocities of individual vehicles recorded by the detector is

$$V_\ell = \{v_{k\ell} \in \mathbb{R}^+ | k = 0, 1, 2, \ldots, N\}, \qquad (\ell \in \{0, 1, 2\}). \tag{2}$$

The lengths of vehicles are denoted as $d_{k\ell}$ and summarized in the set D_ℓ. We remark that the indices $\ell = 0, 1, 2$ correspond to the slow, main and overtaking lanes of a freeway, respectively. The slow lane is intended predominantly for long vehicles (trucks or lorries) and therefore the respective data will not be considered for this research. On the contrary, the drivers use the third lane ($\ell = 2$) particularly if they are overtaking (and after an overtaking-manoeuvre they turn back in the main lane), or if they are significantly faster than other cars. All of the above-mentioned quantities are (for the purposes of this paper) considered to be primary ones, which means that they are directly detectable by the traffic detectors.

Except primary traffic quantities, we now introduce some important secondary quantities, whose values are obtained vicariously, i.e. they are not included in the traffic detector's records. The cardinal secondary traffic micro-quantities are *the time-headways*

$$z_{k\ell} := \tau_{k\ell}^{(\mathrm{in})} - \tau_{(k-1),\ell}^{(\mathrm{in})} \tag{3}$$

and *time clearances*

$$t_{k\ell} := \tau_{k\ell}^{(\mathrm{in})} - \tau_{(k-1),\ell}^{(\mathrm{out})}. \tag{4}$$

Although both of them are not explicitly included in traffic-data files, their values are not burdened with any additional error. Above that, brutto space-gaps between successive vehicles (usually called *distance-headways*) are traditionally approximated by the relation $s_{k\ell} := v_{k\ell} z_{k\ell}$ that presupposes the constant velocity $v(\tau) = v_{k\ell}$ during a time period when $\tau \in [\tau_{(k-1),\ell}^{(\mathrm{in})}, \tau_{k\ell}^{(\mathrm{in})}]$. As is well known, such a precondition is questionable, especially in the region of small traffic densities where the time-headways are too large. However, the influence of a possible error is of marginal importance, as is apparent from the fact that the headway distributions analyzed in small-density regions do not show any noticeable deviation from

the exponential distribution expected for infrequent events (see [2, 13], or [6]). Similarly, *the distance-clearance* is calculated via $r_{k\ell} := v_{k\ell}t_{k\ell}$ and represents the estimated netto distance between the kth car and its predecessor. Contrary to time-headways (clearances), distance-headways (clearances) are burdened by the systematic error that is discussed above. Such an error devalues the knowledge on distance-headway distributions and their evolution.

Denoting the sampling size by m (in this study there is consistently considered $m = 50$) and the number of data sets by M_ℓ (which therefore implies that $M_\ell = \lfloor N_\ell/m \rfloor$, where $N_\ell = \left|\left\{\tau_{k\ell}^{(\mathrm{in})} | \ell \text{ is fixed}\right\}\right|$), one acquires the main data samples

$$S_j^{(\mathrm{main})} = \left\{\left(\tau_{k\ell}^{(\mathrm{in})}, \tau_{k\ell}^{(\mathrm{out})}, v_{k\ell}, d_{k\ell}\right) \in T_1^{(\mathrm{in})} \times T_1^{(\mathrm{out})} \times V_1 \times D_1 | \right.$$
$$\left. k = (j-1)m+1, (j-1)m+2, \ldots, jm \wedge \ell = 1\right\}, \tag{5}$$

and the secondary data samples

$$S_i^{(\mathrm{fast})} = \left\{\left(\tau_{k\ell}^{(\mathrm{in})}, \tau_{k\ell}^{(\mathrm{out})}, v_{k\ell}, d_{k\ell}\right) \in T_2^{(\mathrm{in})} \times T_2^{(\mathrm{out})} \times V_2 \times D_2 | \right.$$
$$\left. k = (i-1)m+1, (i-1)m+2, \ldots, im \wedge \ell = 2\right\}, \tag{6}$$

where $j = 1, 2, \ldots, M_1$ and $i = 1, 2, \ldots, M_2$. For each data sample $S_j^{(\mathrm{main})}$ (or alternatively $S_i^{(\mathrm{fast})}$) we calculate the local flux

$$J_j = \frac{m}{\tau_{jm,\ell}^{(\mathrm{out})} - \tau_{(j-1)m+1,\ell}^{(\mathrm{in})}} \tag{7}$$

and local average velocity $\bar{v}_j = m^{-1} \sum_{k=(j-1)m+1}^{jm} v_{k\ell}$. The local density ϱ_j is then estimated (as suggested in [2]) via the fluid-dynamic equation

$$\varrho_j = \frac{J_j}{\bar{v}_j}. \tag{8}$$

Such a expression is understood as one of the approximations suitable for the estimation of vehicular density. We add that the incorrectness of the definition (8) is caused by the mixing of time and spatial averaging. Besides the macroscopic description of each sample S_j, we now introduce the mean time clearance (or distance-clearance) by means of definitions

$$\bar{t}_j = \frac{1}{m} \sum_{k=(j-1)m+1}^{jm} t_{k\ell}, \qquad \bar{r}_j = \frac{1}{m} \sum_{k=(j-1)m+1}^{jm} r_{k\ell}, \tag{9}$$

respectively.

Since aiming to investigate the essential properties of micro-distributions (i.e. statistical distributions of microscopic traffic quantities) we eliminate (in the next part of this text) the global trends in those distributions. Predominantly, we eliminate the changes of the average headways caused by the varying traffic density. Such an approach corresponds to the technique examined in [8, 11, 13, 26], or [19] and allows more sophisticated comparisons for different traffic regimes or for traffic data originating from different countries. The procedure of the headway re-scaling (see the text below) represents in fact the trivial variant of the so-called *Savitzky–Golay smoothing filter* (for details see [29]) applied to the matrix spectra in the Random Matrix Theory (see [30]), for example. Thus, we define the *scaled distance-clearances* (for the sample S_j) as

$$x_{k\ell} = \frac{r_{k\ell}}{\bar{r}_j} \quad \text{for all} \quad k \in \{(j-1)m+1, (j-1)m+2, \ldots, jm\} \tag{10}$$

and the *modified velocities* (representing a local traffic flow, in fact) as

$$w_{k\ell} = \frac{v_{k\ell}}{\bar{r}_j} \quad \text{for all} \quad k \in \{(j-1)m+1, (j-1)m+2, \ldots, jm\}. \tag{11}$$

It implies that the mean clearance in each sample is re-scaled to the unit. In figure 1 we demonstrate the changes of averages of the modified velocities $w_{k\ell}$. As is understandable, such a dependence corresponds to the shape of the fundamental diagram.

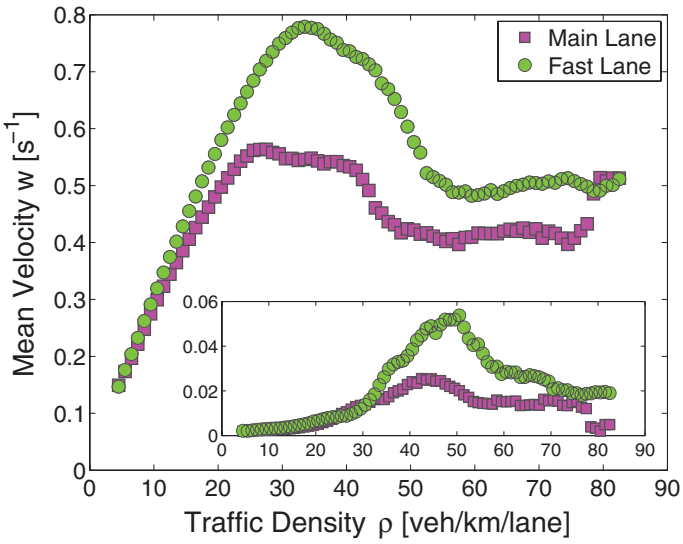

Figure 1. Mean modified velocity w (and variance) as a function of traffic density. The main part of the figure displays the average value $\bar{w}$, which is drawn separately for the fast lane and the main lane. The fluctuations of velocity quantified by the velocity-variance $\mathrm{VAR}(W) = n^{-1}\sum_{k=1}^{n}(w_k - \bar{w})^2$ depending on traffic density, are visualized in the inset. In fact, this diagram corresponds to a fundamental diagram in the re-scaled form.

3. Distance correlations among individual traffic data

For intentions of analytical calculations executed in section 5, it is now necessary to test the statistical dependence between individual traffic data. Although the classical way of inspecting this property is to calculate the Pearson's correlation coefficient, in our paper we will consistently use the advanced approaches discussed in [20–22]. The reason for such a novel approach is that the Pearson's correlation coefficient (contrary to the distance correlation introduced below) does not reveal nonlinear or non-monotone dependences in general.

Thus, consider two random vectors Y and Z with finite first moments. Let $f_Y(y)$, $f_Z(z)$ represent the associated probability densities respectively, whereas $f_{YZ}(y, z)$ represents the joint probability density. Then the *distance-covariance* $\mathrm{COV}_d(Y, Z)$ between Y and Z is defined by

$$\mathrm{COV}_d^2(Y,Z) = \frac{1}{\mathrm{C}_p\mathrm{C}_q}\int_{\mathbb{R}^{p+q}} \frac{|f_{YZ}(y,z) - f_Y(y)f_Z(z)|^2}{\|y\|_p^{p+1}\|z\|_q^{q+1}}\mathrm{d}y\,\mathrm{d}z, \tag{12}$$

where

$$\mathrm{C}_r = \frac{\pi^{(r+1)/2}}{\Gamma[(r+1)/2]} \tag{13}$$

and $\|x\|_r$ is the Euclidean norm of x in r-dimensional space $\mathbb{R}^r$. Using this notation the *distance correlation* ω_{yz} between Y and Z is then defined by the relation

$$\omega_{yz}^2 = \begin{cases} \frac{\mathrm{COV}_d^2(Y,Z)}{\sqrt{\mathrm{COV}_d^2(Y,Y)\cdot\mathrm{COV}_d^2(Z,Z)}} & \dots \quad \mathrm{COV}_d^2(Y,Y)\cdot\mathrm{COV}_d^2(Z,Z) > 0, \\ 0 & \dots \quad \mathrm{COV}_d^2(Y,Y)\cdot\mathrm{COV}_d^2(Z,Z) = 0. \end{cases} \tag{14}$$

The above-defined distance-covariance and distant-correlation are analogous to the standard covariance and correlation, but they generalize and extend these classical bivariate measures in the following sense. The distance correlation ω (contrary to Pearson's correlation coefficient) is zero if and only if the random vectors are independent, which is mathematically

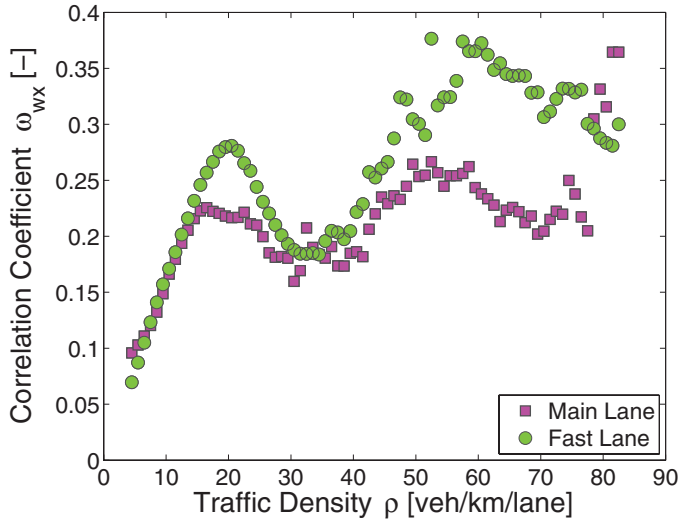

Figure 2. Distance-correlation coefficient ω_{wx} as a function of traffic density. The value ω_{wx} quantifies the distance correlation (14) between the modified vehicular velocities $w_{k\ell}$ and scaled gaps $x_{k\ell}$ to the preceding car (in the regions of fixed traffic densities). The squares represent cars moving in the main lane, whereas circles correspond to the fast-lane cars. Note that the vehicles in the fast lane show the stronger correlations. Such a tendency is accented in the regions where the mental strain of drivers is stronger (i.e. in the regions of over-saturations).

profitable. Another pleasant property of ω, as well as the method for calculating the sample distance-covariance, are summarized in [22].

First of all, we apply the presented procedure to the sets $X = \{x_1, x_2, \ldots, x_n\}$ and $W = \{w_1, w_2, \ldots, w_n\}$ of scaled distance-clearances and modified velocities. Dividing the complete data sample into the sub-samples of almost-constant traffic density we demonstrate in figure 2 that coefficient ω_{wx} is localized in the range $[0.05; 0.40]$ and the corresponding values depend on a traffic phase and freeway lane. Although the investigated correlation is restrained, it is evident that individual clearances and velocities are correlated, which was expected as a consequence of platoon formation (see [12]).

4. Theoretical scheme of vehicular model used

As introduced in [8], justified in [16] and elaborated in [13], one of the possible methods of acquiring theoretical predictions for distributions of empirical distance-clearances is to use the socio-physical traffic model, whose elements are repulsed by the short-ranged power-law forces and randomized by the socio-physical noise, which is supposed to be of a thermal-like nature. The influence of such a thermodynamical component to the model can be increased/reduced by the socio-physical coefficient $\beta \in [0, \infty)$ reflecting the mental strain which the drivers are under during a given traffic situation. Specifically, in free-traffic regimes (where the psychological pressure caused by the traffic situation is weak) the parameter β is almost zero. On the contrary, for the congested phase (where interactions among drivers are reinforced) the mental strain coefficient β is large. For the detailed changes of β one can inspect [8, 13], or [26].

To be specific, we consider dimensionless particles moving along a ring whose velocities are $w_1, w_2, \ldots, w_N$ and mutual distances (between subsequent particles) are $x_1, x_2, \ldots, x_N$. Introducing the short-ranged repulsive potential

$$U(x_1, x_2, \ldots, x_N) = \sum_{k=1}^{N} \frac{1}{x_k} \tag{15}$$

and socio-physical Hamiltonian (see [13])

$$\mathcal{H}(w_1, w_2, \ldots, w_N, x_1, x_2, \ldots, x_N) = \frac{1}{2}\sum_{k=1}^{N}(w_k - w_\mathrm{d})^2 + U(x_1, x_2, \ldots, x_N) \tag{16}$$

one can derive (for details please see [13]) that velocities of such an ensemble (analyzed in the steady state) are Gaussian-distributed, i.e. the related probability density reads

$$q(w) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(w-w_\mathrm{d})^2}{2\sigma^2}}, \tag{17}$$

where σ^2 is the second statistical moment (variance) and w_d is the optimal velocity of drivers. By analogy, in [13, 26] it can be deduced that the scaled distance-clearance distribution reads as

$$\wp(r) = A\Theta(r)\, \mathrm{e}^{-\beta/r}\, \mathrm{e}^{-Dr}, \tag{18}$$

where $\Theta(x)$ is the Heaviside's step-function and the normalization factors are

$$D \approx \beta + \frac{3 - \mathrm{e}^{-\sqrt{\beta}}}{2}, \tag{19}$$

$$A^{-1} = 2\sqrt{\frac{\beta}{D}}\mathcal{K}_1(2\sqrt{D\beta}). \tag{20}$$

As verified in [8, 13, 24] the one-parametric distribution-family (18) is in satisfactory agreement with the clearance distribution observed in real-road data. We add that the noise-parameter β is related to the corresponding traffic density ϱ.

As is evident, the delineated model represents only an idealized representation of real-road traffic, formulated with respect to its analytical solvability. In fact, such a scheme infringes several basic principles that are generally attributed to vehicular traffic. (1) interaction among cars is not symmetrical. (2) Vehicular traffic is regarded as a system driven far from equilibrium. (3) Vehicular interactions have only a finite range. These discrepancies can be, however, eliminated if formulating a conveniently-adjusted numerical representation of the model. Indeed, numerical schemes (based on the adaptive Verlet algorithm [6, 33] or on principles similar to those used during the so-called simulated annealing [34]) working with a forwardly-directed power-law and finite-ranged interaction-forces produce typical effects for realistic vehicular trajectories (see [31, 35]). On the other hand, associated distributions of micro-quantities show practically no conspicuous deflections from the ideal variant of the thermodynamical gas introduced. Furthermore, such a 'more realistic version' of the model produces headway distributions of type (18), even if observed far from the thermal balance. Simple numerical experiments confirm that an ensemble of particles exposed to the influence of external random forces (quantified by noise-coefficient ν) passes through states, whose headway distributions correspond to the formula (18) for a fitted value of β_fit. In addition to that, the equality $\beta_\mathrm{fit} \approx \nu$ holds only for states near the equilibrium (if reached). Such changes of the velocity/gap distributions are in agreement with those detected in vehicular flows.

Moreover, in [16], the authors compared two different ensembles with: (1) symmetrical interaction-forces fulfilling Newton's action–reaction law and (2) asymmetric, forwardly-directed interactions and driving terms. They found that the steady-state velocity and distance distributions for asymmetric interactions agree with the equilibrium distributions of classical many-particle systems with symmetrical interactions, provided the system is large enough.

Their result establishes the possibility of approximating the steady-state statistics in driven many-particle systems by Hamiltonian systems, which fully corresponds with outputs of experiments described in the previous paragraph.

5. Mathematical derivation of time-clearance distribution

With the help of formulas (17) and (18) one can mathematically derive the analytical form of probability density $\eta(t)$ for clear time intervals between two consecutive vehicles (particles of a model). Since the joint probability density for distance and velocity can be (under the condition on weak distance-versus-speed dependence) predicted as $g(w,x) = q(w)\wp(x)$, the joint probability density for time gaps and velocities therefore reads as $h(w,t) = wq(w)\wp(wt)$. Then the time-clearance distribution represents a marginal density

$$\eta(t) = \int_{\mathbb{R}} h(w,t)\,\mathrm{d}w = \int_{\mathbb{R}} wq(w)\wp(wt)\,\mathrm{d}w. \tag{21}$$

After expanding a function $f(w) = w\wp(wt)$ into the Taylor's series about the optimal velocity w_d we acquire

$$f(w) = f(w_\mathrm{d}) + \sum_{\ell=1}^{\infty} \frac{1}{\ell!}\frac{\mathrm{d}^\ell f}{\mathrm{d}w^\ell}(w_\mathrm{d})(w-w_\mathrm{d})^\ell, \tag{22}$$

where

$$\frac{\mathrm{d}^\ell f}{\mathrm{d}w^\ell} = \frac{\partial^\ell \wp}{\partial(wt)^\ell}t^\ell w + \ell\frac{\partial^{\ell-1}\wp}{\partial(wt)^{\ell-1}}t^{\ell-1} \quad (\ell \in \mathbb{N}). \tag{23}$$

Hence

$$\begin{aligned}\eta(t) &= w_\mathrm{d}\wp(w_\mathrm{d}t) + \sum_{\ell=1}^{\infty}\frac{1}{\ell!}\frac{\partial^\ell\wp}{\partial(wt)^\ell}(w_\mathrm{d}t)t^\ell w_\mathrm{d}\int_{\mathbb{R}}(w-w_\mathrm{d})^\ell q(v)\,\mathrm{d}w \\ &\quad + \sum_{\ell=1}^{\infty}\frac{1}{\ell!}\frac{\partial^{\ell-1}\wp}{\partial(wt)^{\ell-1}}(w_\mathrm{d}t)t^{\ell-1}\int_{\mathbb{R}}(w-w_\mathrm{d})^\ell q(w)\,\mathrm{d}w \\ &= w_\mathrm{d}\wp(w_\mathrm{d}t) + \sum_{\ell=1}^{\infty}\frac{\mu_\ell}{\ell!}\left(\frac{\partial^\ell\wp}{\partial(wt)^\ell}(w_\mathrm{d}t)t^\ell w_\mathrm{d} + \frac{\partial^{\ell-1}\wp}{\partial(wt)^{\ell-1}}(w_\mathrm{d}t)t^{\ell-1}\right),\end{aligned} \tag{24}$$

where $\mu_\ell = \int_{\mathbb{R}}(w-w_\mathrm{d})^\ell q(w)\,\mathrm{d}w$ is ℓth central statistical moment with three prerogative cases $\mu_0 = 1$, $\mu_1 = 0$, and $\mu_2 = \sigma^2$. The latter represents a statistical variance (see also (17)). As is well known, the odd central statistical moments (related to the Gaussian distribution) are zero and even moments comply with equalities $\mu_{2\ell} = \sigma^{2\ell}(2\ell-1)!!$, which leads to the general formula

$$\eta(t) = w_\mathrm{d}\wp(w_\mathrm{d}t) + \sum_{\ell=1}^{\infty}\frac{\sigma^{2\ell}}{(2\ell)!!}\left(\frac{\partial^{2\ell}\wp}{\partial(wt)^{2\ell}}(w_\mathrm{d}t)t^{2\ell}w_\mathrm{d} + \frac{\partial^{2\ell-1}\wp}{\partial(wt)^{2\ell-1}}(w_\mathrm{d}t)t^{2\ell-1}\right). \tag{25}$$

Owing to the facts $\int_{\mathbb{R}} w_\mathrm{d}\wp(w_\mathrm{d}t)\,\mathrm{d}t = 1$ and

$$\int_{\mathbb{R}}\left(\frac{\partial^{2\ell}\wp}{\partial(wt)^{2\ell}}(w_\mathrm{d}t)t^{2\ell}w_\mathrm{d} + \frac{\partial^{2\ell-1}\wp}{\partial(wt)^{2\ell-1}}(w_\mathrm{d}t)t^{2\ell-1}\right)\mathrm{d}t = 0, \tag{26}$$

the TC-distribution (25) is normalized correctly. In a certain sense (as is understandable from the two previous equations) the first factor in (25) forms a leading term of the TC-distribution, whereas the sum in (25) represents a perturbation term only. Above that, the weight of the ℓth perturbation summands decreases with a factor $\sigma^{2\ell}/(2\ell)!!$. For practical applications

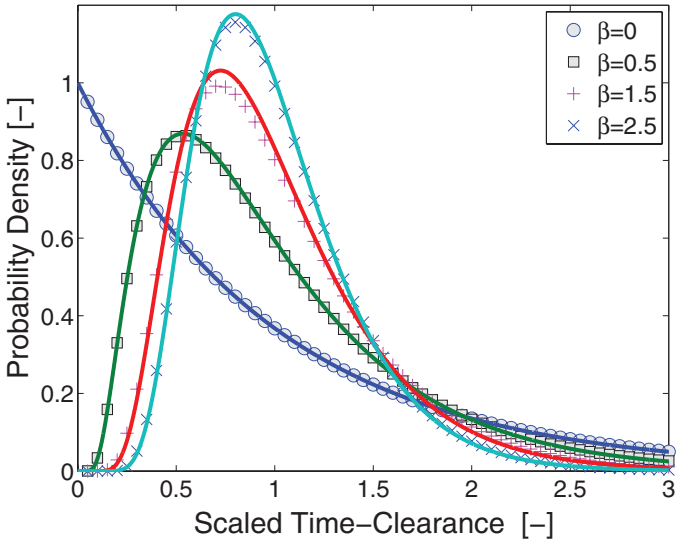

Figure 3. Time-clearance distribution (analytically derived). The signs represent the values of function (28) calculated for $\beta \in \{0, 1/2, 3/2, 5/2\}$, whereas the curves display the zero approximation (29) for the same β. Small discrepancies between signs and corresponding curves are caused by the fact that statistical variance σ^2 of the modified velocities (measured on freeways) is very low. Phenomenologically, this relation can be estimated (for the main-lane data) by the inequality $\sigma^2 \leqslant (45\beta/55)^4 \exp[-45\beta/14]$.

(especially for traffic applications), it seems to be coherent, therefore, to use an approximate expansion

$$\eta(t) \approx w_{\mathrm{d}}\wp(w_{\mathrm{d}}t) + \frac{\sigma^2}{2}\left(\frac{\partial^2\wp}{\partial(wt)^2}(w_{\mathrm{d}}t)t^2w_{\mathrm{d}} + 2\frac{\partial\wp}{\partial(wt)}(w_{\mathrm{d}}t)t\right). \tag{27}$$

Such an approximation is legitimate because of the extremely low value of velocity variance (see the inset in figure 1). Note that the factors $\sigma^{2\ell}/(2\ell)!!$ measured in real-road data rapidly vanish since the maximal value of the velocity variance is $\approx$0.02 in the main lane or $\approx$0.06 in the fast lane. The expansion (27) can be consecutively replaced (after the scaling procedure) by the formula

$$\eta(t) \approx \wp(t) + \sigma^2\wp(t)\left(\frac{\beta^2}{2t^2} + \frac{D^2t^2}{2} - D(t+\beta)\right). \tag{28}$$

The influence of the low velocity variance to the TC-distribution is demonstrated in figure 3 where the curves (28) are compared to the zero approximation

$$\eta(t) \approx A\Theta(t)\,\mathrm{e}^{-\frac{\beta}{t}}\,\mathrm{e}^{-Dt}, \tag{29}$$

while the relations (19), (20) hold true. We remind the reader that the normalization and re-scaling conditions $\int_{\mathbb{R}}\eta(t)\,\mathrm{d}t = \int_{\mathbb{R}} t\,\eta(t)\,\mathrm{d}t = 1$ have been applied.

6. Criteria for acceptability of analytical clearance distributions

Owing to the empirical background of the topic investigated in this research, the curves representing theoretical approximations of general clearance distributions (we denote them by $g(s)$, in general) have to fulfil certain mathematical and empirically-inspired criteria. Except the pure mathematical properties of $g(s)$ (non-negativity, normalization, positivity of

the support, scaling property and continuity), some other requirements can be derived from recent knowledge on the microscopic structure of vehicular samples. As is apparent from many scientific sources (see [2, 5, 17, 25–28, 32]), the spatial or temporal headway/clearance distributions (analyzed for congested traffic streams) show the heavy plateau located near the origin. Such a plateau can be understood as a consequence of the strong repulsions among closely-occurring vehicles whose drivers make an effort to prevent a possible crash. Mathematically, such a phenomenon is described by means of the following definition:

$$\forall m \in \mathbb{N} : \lim_{s \to 0_+} s^{-m} g(s) = 0, \tag{30}$$

which is (for the locally smooth densities $g(s) \in \mathscr{C}^{\infty}(0, \delta)$) equivalent to the conditions $\frac{\mathrm{d}^m g}{\mathrm{d}s^m}(0_+) = 0$. Unfortunately, $g(s)$ is not (as immediately follows from the previous lemma) an analytical function.

The second empirically-inspired criterion is induced from the perspicuous fact that all vehicular interactions are finite-ranged, i.e., the movements of two sufficiently outlying cars are not correlated (even in congested traffic). Such a statistical ensemble is usually entitled as *quasi-Poissonian.* This terminology reflects the well-known fact that a system is qualified as *Poissonian (purely Poissonian)* if all related subsystems are independent. In this case, the probability for the occurrence of several elements inside the fixed (space or time) region conforms to the Poissonian distribution. If interactions among elements are restricted to several neighbors only, the Poissonian features of adjacent elements are broken. On the contrary, outlying elements still behave as independent, which leads to the similarity between the tailed distribution and that derived for pure Poissonian ensembles. Therefore, the tails of headway distributions (for pure and quasi-Poissonian ensembles, respectively) show similar trends. These considerations result in the undermentioned definition.

Definition 6.1. *A probability density $g(s)$ (and the associated distribution function) is called balanced if there exists $\omega > 0$ so that*

$$\forall \varkappa > \omega : \quad \lim_{s \to +\infty} g(s)\, \mathrm{e}^{\varkappa s} = +\infty, \tag{31}$$

and

$$\forall \varkappa \in (0, \omega) : \quad \lim_{s \to +\infty} g(s)\, \mathrm{e}^{\varkappa s} = 0. \tag{32}$$

The number ω is then called the balancing index and denoted by $\mathrm{inb}(g)$. *The class of balanced distributions is denoted by $\mathscr{B}$.*

As is apparent, the family of balanced distributions and the family of heavy-tailed distributions (see [36, 37]) are disjointed. Thus, the intersection of $\mathscr{B}$ and the class of subexponential distributions is empty. The same also holds true for the class of fat-tailed or long-tailed distributions. Therefore, the class $\mathscr{B}$ is a special subclass of light-tailed distributions. To conclude, based on an assumption that vehicular interactions are finite-ranged, all associated gap distributions should also meet the so-called *balance criterion:* $g(s) \in \mathscr{B}$.

In conclusion, we point out that the distributions (28) and (29) meet all the above-cited criteria and represent, therefore, two appropriate candidates for vehicular headway distribution.

7. Mathematical derivation of multi-clearance distribution

Now, knowing the one parameter family of the probability densities (28) or their zero approximations (29) one can derive an analytical prediction for the so-called *nth multi-clearance distribution* $\tau_n(t)$ which represents the probability density for a clear time gap t among $n + 2$ neighboring particles. Therefore, the main goal of this section is to quantify (by

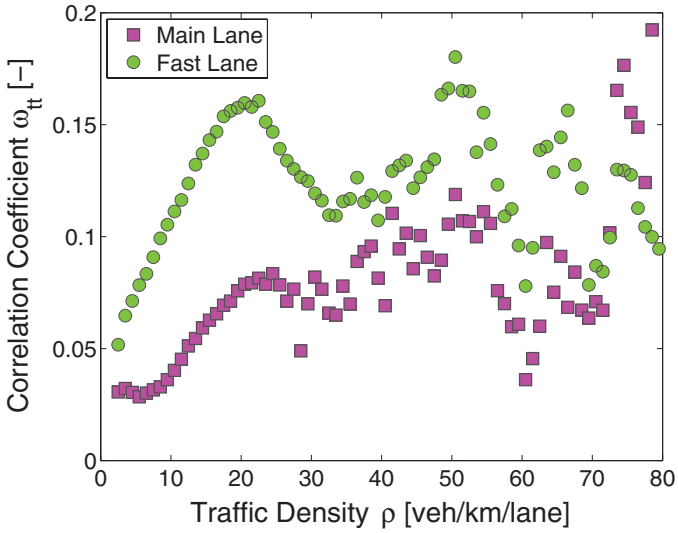

Figure 4. Graphical representation of distance-correlation analysis of successive time clearances. The values $\omega_{tt} = \omega_{tt}(\varrho)$ describe the distance-correlation coefficient (14) calculated for pairs of succeeding time clearances extracted from traffic data in various density regions.

means of a theory of functional convolutions) the probability that the time period between two following instants (the first one: the back bumper of the kth vehicle is leaving the detector; the second one: the front bumper of the $(k+n+1)$th vehicle is intersecting the detector line) ranges in the interval $[t, t+\mathrm{d}t)$. Using this notation, we find that the probability density for the time clearance between two succeeding cars (derived in the previous section) is $\eta(t) = \eta_0(t)$. Regarding the clearances as independent (which is here understood as a theoretical surmise—see figure 4) the nth probability density $\eta_n(t)$ can be calculated via recurrent formula

$$\eta_n(t) = \eta_{n-1}(t) \star \eta_0(t), \tag{33}$$

where symbol $\star$ represents a convolution of the two probabilities, i.e.,

$$\eta_n(t) = \int_{\mathbb{R}} \eta_{n-1}(s)\eta_0(t-s)\,\mathrm{d}s. \tag{34}$$

As a first resort, we will concentrate our endeavor to studying the low-density traffic regimes where the locations of cars are independent and therefore the traffic flow corresponds (from a mathematical point of view) to the Poisson process of uncorrelated events. As published in [8, 13, 26], such a situation is characterized by the negligible value of the socio-physical coefficient β reflecting a driver's mental strain. Thus, the related clearance distribution reads $\eta(t) = \Theta(t)\mathrm{e}^{-t}$ and sequentially the multi-clearance distribution for free traffic looks like

$$\eta_n(t) = \Theta(t)\,\mathrm{e}^{-t} \int_0^t \int_0^{s_1} \dots \int_0^{s_{n-1}} \mathrm{d}s_n \dots \mathrm{d}s_1 = \Theta(t)\frac{t^n}{n!}\mathrm{e}^{-t}. \tag{35}$$

As a second resort, we aim to derive a general formula, i.e. a formula for non-zero β. For these purposes, we use the zero approximation (29). Applying the method of mathematical induction and an approximation of the function

$$g_n(t,s) = \mathrm{e}^{-\beta\left(\frac{n^2}{s}+\frac{1}{t-s}\right)} \approx \mathrm{e}^{-\frac{\beta}{t}(n+1)^2} \tag{36}$$

in the saddle point one can obtain

$$\begin{aligned}\eta_n(t) &= \Theta(t)\int_0^t A_{n-1}As^{n-1}\mathrm{e}^{-\beta\frac{n^2}{s}}\mathrm{e}^{-Ds}\mathrm{e}^{-\frac{\beta}{t-s}}\mathrm{e}^{-D(t-s)}\,\mathrm{d}s\\ &= \Theta(t)A_{n-1}A\mathrm{e}^{-Dt}\int_0^t s^{n-1}g_n(t,s)\,\mathrm{d}s\\ &\approx \Theta(t)A_{n-1}A\mathrm{e}^{-\frac{\beta}{t}(n+1)^2}\mathrm{e}^{-Dt}\int_0^t s^{n-1}\,\mathrm{d}s\\ &\approx \Theta(t)A_{n-1}An^{-1}t^n\mathrm{e}^{-\frac{\beta}{t}(n+1)^2}\mathrm{e}^{-Dt}. \end{aligned} \tag{37}$$

Hence

$$\eta_n(t) \approx \Theta(t)A_n t^n\mathrm{e}^{-\frac{\beta}{t}(n+1)^2}\mathrm{e}^{-Dt}, \tag{38}$$

where (after applying the re-normalization procedure)

$$A_n^{-1} = 2\left(\sqrt{\frac{\beta}{D}}(n+1)\right)^{n+1}\mathcal{K}_{n+1}(2(n+1)\sqrt{D\beta}). \tag{39}$$

This fixes the proper normalization $\int_{\mathbb{R}}\eta_n(t)\,\mathrm{d}t = 1$. In addition to that the mean nth spacing equals to

$$\int_{\mathbb{R}} t\eta_n(t)\,\mathrm{d}t = n+1. \tag{40}$$

Note, that (38) holds true also for the limiting case $\beta = 0$. Really, with help of formulas $\lim_{x\to 0}x^{n+1}\mathcal{K}_{n+1}(x) = (2n)!!$ and $\lim_{\beta\to 0_+}D(\beta) = 1$ we easily deduce that

$$\lim_{\beta\to 0_+} 2\left(\sqrt{\frac{\beta}{D}}(n+1)\right)^{n+1}\mathcal{K}_{n+1}(2(n+1)\sqrt{D\beta}) = n!. \tag{41}$$

Therefore it holds $A_n = 1/n!$, which is in full agreement with the relation (35). Thus, the relation (38) constitutes a zero approximation for the distribution of time multi-clearances. Since we have supposed (in the previous deductions) that the variance of modified velocities is negligible (in a local sense), the form of the distribution (38) is a direct consequence of the distribution for spatial clearances and the fact that all cars have practically the same velocity (in a local sense, again).

Because the empirical measurements show low (but not negligible) variances in modified velocities (see the inset in figure 1), it seems more realistic to derive the time multi-clearance distributions under the conditions $\sigma^2 > 0$ and $\sigma^{2n} \approx 0$ for $n = 2, 3, 4, \ldots$. Thus, we suppose that clear time intervals among succeeding cars are distributed according to the rule (28). The detailed analysis of such a relation vindicates the fact that the dominating term in the last summand of (28) is $\sigma^2 Dt\,\wp(t)$. Hence, we surmise that the time clearance follows the law

$$\eta_0^{(\sigma)}(t) \approx \wp(t) - \sigma^2 Dt\,\wp(t). \tag{42}$$

Here we remark (for mathematical correctness) that this function has to be understood as an approximative probability density, since the perturbation term $\sigma^2 Dt\,\wp(t)$ causes that $\eta_0^{(\sigma)}(t)$ does not fulfil the exact mathematical definition.

Using the method of mathematical induction, we prove below that (under the previous surmises) the multi-clearance distribution of order n reads

$$\eta_n^{(\sigma)}(t) \approx \Theta(t)\frac{A^{n+1}}{n!}t^n\mathrm{e}^{-(n+1)^2\frac{\beta}{t}}\mathrm{e}^{-Dt}(1-\sigma^2 Dt). \tag{43}$$

For completeness, we remark that in the following mathematical operations

$$\begin{aligned}\eta_{n+1}^{(\sigma)}(t) = \eta_n^{(\sigma)}(t) \star \eta_0^{(\sigma)}(t) \approx \Theta(t)\frac{A^{n+2}}{n!}\mathrm{e}^{-Dt}\int_0^t \mathrm{e}^{-(n+1)^2\frac{\beta}{s}} s^n \mathrm{e}^{-\frac{\beta}{t-s}}\,\mathrm{d}s \\ +D^2\sigma^4\Theta(t)\frac{A^{n+2}}{n!}\mathrm{e}^{-Dt}\int_0^t \mathrm{e}^{-(n+1)^2\frac{\beta}{s}}(t-s)s^{n+1}\mathrm{e}^{-\frac{\beta}{t-s}}\,\mathrm{d}s \\ -D\sigma^2\Theta(t)\frac{A^{n+2}}{n!}\mathrm{e}^{-Dt}\int_0^t \mathrm{e}^{-(n+1)^2\frac{\beta}{s}} s^{n+1}\mathrm{e}^{-\frac{\beta}{t-s}}\,\mathrm{d}s \\ -D\sigma^2\Theta(t)\frac{A^{n+2}}{n!}\mathrm{e}^{-Dt}\int_0^t \mathrm{e}^{-(n+1)^2\frac{\beta}{s}}(t-s)s^{n}\mathrm{e}^{-\frac{\beta}{t-s}}\,\mathrm{d}s \\ = \Theta(t)\frac{A^{n+2}}{(n+1)!}t^{n+1}\mathrm{e}^{-(n+2)^2\frac{\beta}{t}}\mathrm{e}^{-Dt}(1-\sigma^2 Dt) \qquad (44)\end{aligned}$$

we have used the approximation (36) and the surmise $\sigma^4 \approx 0$.

8. Empirical multi-clearance distribution versus analytical prediction

In this section we will balance the theoretical prognoses deduced in sections 5 and 7 against the empirical multi-clearance distributions of freeway data analyzed in section 2 (two-lane freeway A9—Netherlands, two-lane freeway D1—Czech Republic). In the following part of the text, we will consider the multi-clearances

$$\varkappa_i^{(j)} = \sum_{k=(j-1)m+i}^{(j-1)m+i+n} \frac{t_{k\ell}}{\bar{t}_j}, \quad (\ell = 1,\, j = 1, 2, \ldots, M_1,\, i = 1, 2, \ldots, m-n) \qquad (45)$$

enumerated for data sets (5). These multi-clearances are associated with the specific traffic density through the formula (8) and re-scaled so that the average multi-clearance (in the given data samples $S_j^{(\mathrm{main})}$) is equal to $n+1$. Now, the multi-clearances $\varkappa_i^{(j)}$ represent a statistical realization of random variable t considered in the section 7 and can therefore be confronted with the theoretical probabilities.

To be factual, we divide the entire interval of traffic densities into the subintervals $[0, 3)$, $[1, 4)$, $[2, 5)$, and so on and analyze the freeway multi-clearance distribution $p(\varkappa)$ separately in each subinterval. A similar approach has been examined in [8, 11, 13, 26] and reflects the known fact that the headway distributions are substantially influenced by the changing location of the traffic ensemble in the phase diagram. To prevent a mixing of states with different vigilances of car drivers (or with different temperatures of the related heat bath—in the thermodynamic interpretation of vehicular traffic), we inspect the empirical multi-clearances separately in each density interval. Such an approach permits the investigation of the density-dependent evolution of multi-clearance distributions.

For comparison purposes, we now introduce the general statistical distance

$$\chi_n(\alpha_0, \alpha_1, \ldots, \alpha_k) = \frac{\mathrm{e}}{n+1}\int_0^\infty |f_n(t|\alpha_0, \alpha_1, \ldots, \alpha_k) - p(t)|^2 t\mathrm{e}^{-\frac{t}{n+1}}\,\mathrm{d}t \qquad (46)$$

cumulating the weighted deviations between a α-parameterized theoretical prediction $f_n(t|\alpha_0, \alpha_1, \ldots, \alpha_k)$ and empirical frequency $p(t)$. We remark that the weight factor $\phi_n(t) = \frac{t}{n+1}\exp\left[1 - \frac{t}{n+1}\right]$ has been chosen *(a)* to eliminate the influence of long clearances, *(b)* to suppress extremely short clearances and finally *(c)* to increase the influence of clearances being close to the mean value. In addition, $\mathrm{argmax}_{t\geqslant 0}\phi_n(t) = n+1$ and $\phi_n(n+1) = 1$.

The illustrative examples of such a comparison procedure are depicted in figures 5 and 6 where we plot empirical multi-clearance distributions (histograms) against the Poissonian

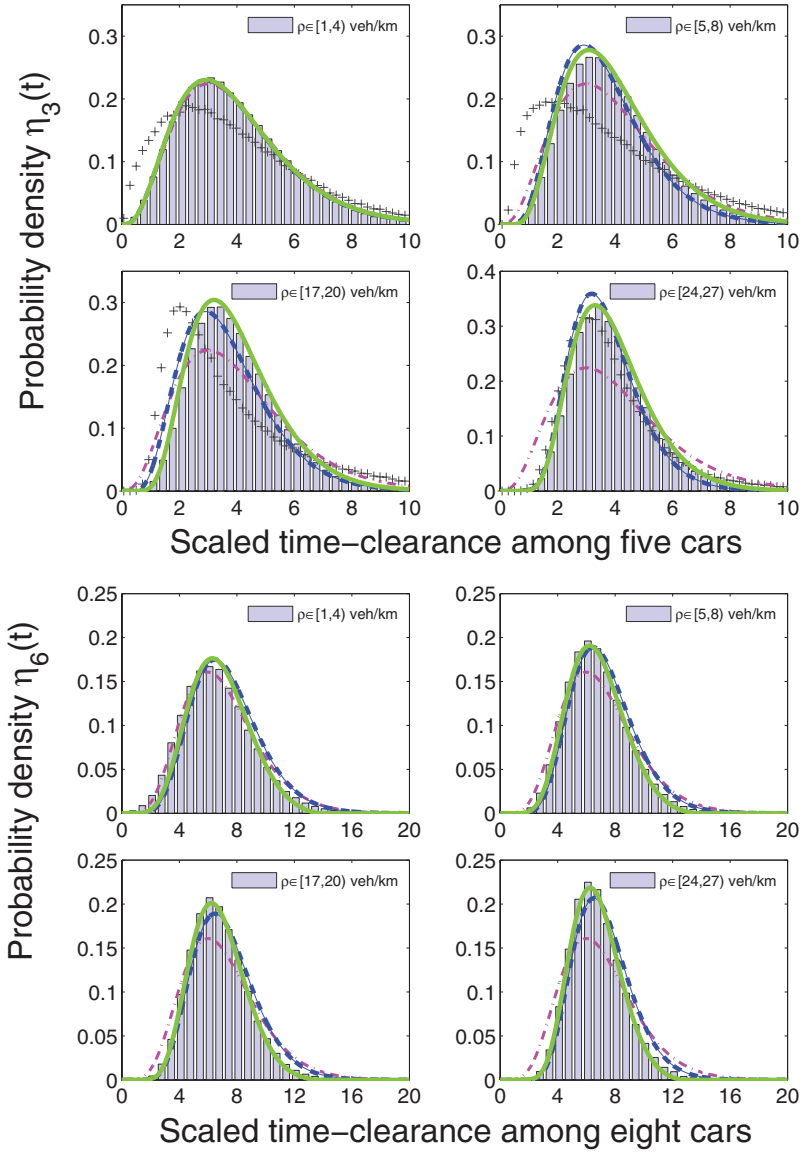

Figure 5. Statistical distribution of time multi-clearances for low-density regimes. The bars shows the empirical probability density for the clear time interval among five (or eight) succeeding vehicles moving in the main lane (for various density regimes—see legend for details). The blue dashed curves represent the prediction (38) plotted for $n = 3$ (top plot) or $n = 6$ (bottom plot) and for the fitted value of coefficient $\beta = \beta_{\text{fit}}$ obtained by minimizing the weighted error-function according to the formula (48). Continuous curves (green) display the analytical approximation (47) plotted for the fitted values of $\beta = \beta_{\text{fit}}$ and $\varepsilon = \varepsilon_{\text{fit}}$ specified by minimizing the weighted error-function according to the formula (49). For clearness, we also plot the dash–dotted curves (magenta) visualizing the multi-clearance distribution (35) valid for an occurrence of independent events. Plus signs illustrate how the time multi-clearances (gauged by the fast-lane-detectors) differ from those detected in the main lane.

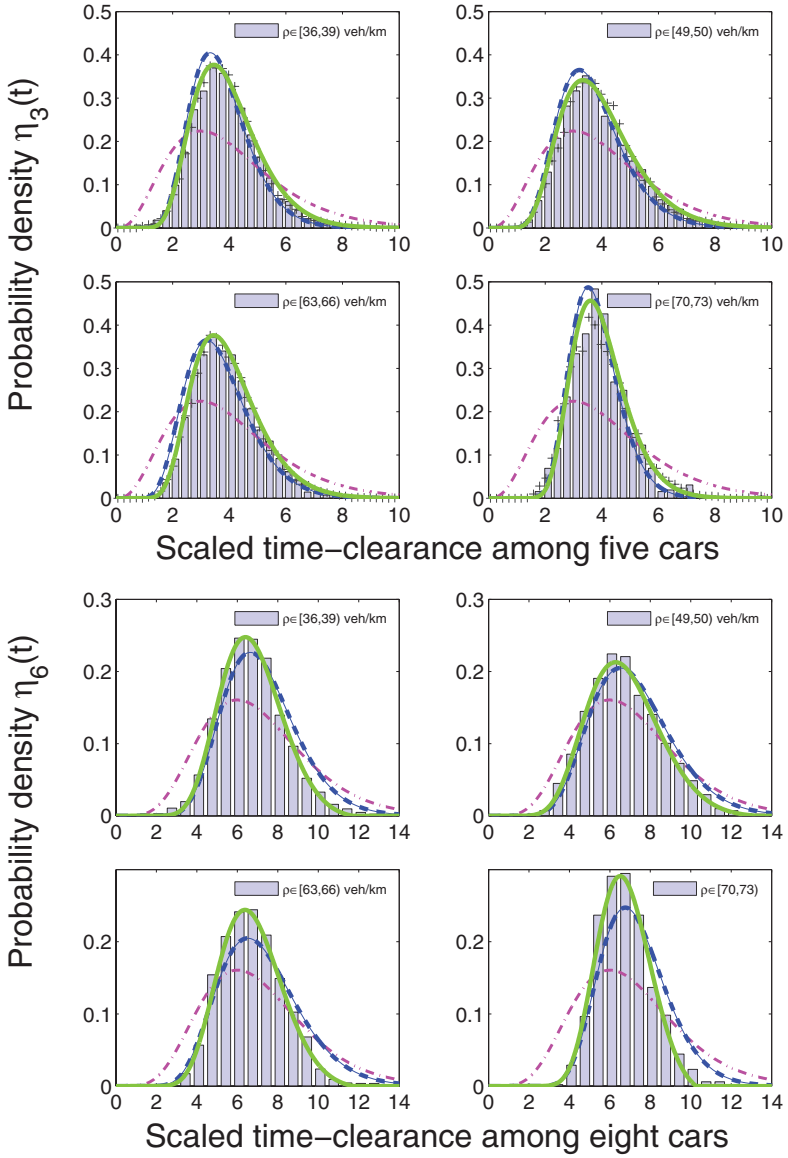

Figure 6. Statistical distribution of time multi-clearances for high-density regimes. The bars show the empirical probability density for a clear time interval among five (or eight) succeeding vehicles moving in the main lane (for various density regimes—see legend for details). The blue dashed curves represent the prediction (38) plotted for $n = 3$ (top plot) or $n = 6$ (bottom plot) and for the fitted value of coefficient $\beta = \beta_{\text{fit}}$ obtained by minimizing the weighted error-function according to the formula (48). Continuous curves (green) display the analytical approximation (47) plotted for the fitted values of $\beta = \beta_{\text{fit}}$ and $\varepsilon = \varepsilon_{\text{fit}}$ specified by minimizing the weighted error-function according to the formula (49). For clearness, we also plot the dash–dotted curves (magenta) visualizing the multi-clearance distribution (35) valid for an occurrence of independent events. Plus signs illustrate how the time multi-clearances (gauged by the fast-lane-detectors) differ from those detected in the main lane.

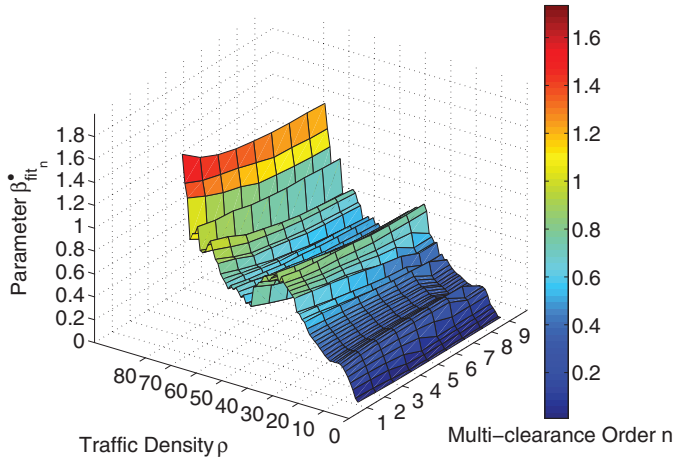

Figure 7. Optimal values of the parameter β as a function of traffic density and multi-clearance order. Colored values help visualize the values of the fitted parameter $\beta^{\bullet}_{\mathrm{fit}_n}$ (see the relation (48) that minimizes the statistical distance between theoretical curve (38) and empirical histogram $p(t)$) enumerated for main-lane data only.

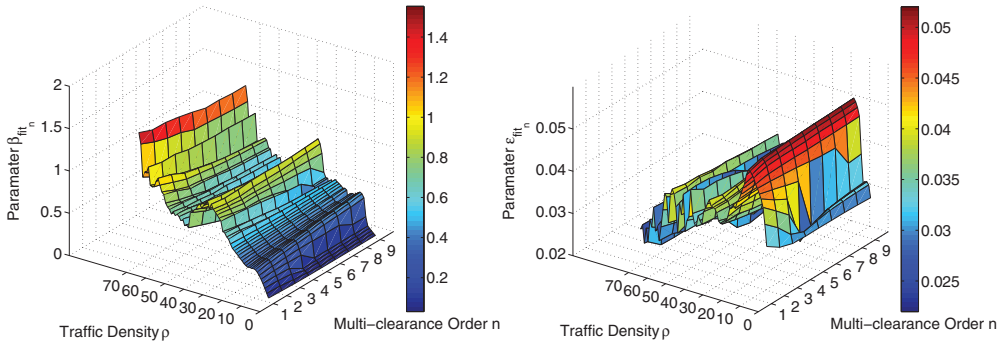

Figure 8. Optimal values of the parameters β and ε in the theoretical distribution (47) as a function of traffic density and multi-clearance order. Colored values visualize the value of the fitted parameter β_{fit_n} and $\varepsilon_{\mathrm{fit}_n}$ (see the relation (49)) that minimizes the statistical distance between the theoretical curve (47) and the empirical histogram $p(t)$ (enumerated for main-lane data only).

distribution (35), zero approximation (38) calculated for $n = 3$ or $n = 6$, and the final analytical prediction

$$\eta_n^{(\varepsilon)}(t) = \Theta(t)\frac{A^{n+1}}{n!}t^n \mathrm{e}^{-(n+1)^2\beta_n/t}\mathrm{e}^{-Dt}(1-\varepsilon_n t). \tag{47}$$

The values of the fitted parameters (β in (38) and β, ε in (47) have been determined by means of formulae

$$\beta^{\bullet}_{\mathrm{fit}_n} = \mathrm{argmin}_{\beta\in[0,\infty)}\int_0^{\infty}\left|\eta_n(t)-p(t)\right|^2 t\,\mathrm{e}^{-\frac{t}{n+1}}\,\mathrm{d}t, \tag{48}$$

$$(\beta_{\mathrm{fit}_n},\varepsilon_{\mathrm{fit}_n}) = \mathrm{argmin}_{\beta\in[0,\infty),\varepsilon\in[0,\infty)}\int_0^{\infty}\left|\eta_n^{(\varepsilon)}(t)-p(t)\right|^2 t\,\mathrm{e}^{-\frac{t}{n+1}}\,\mathrm{d}t. \tag{49}$$

As visible in figures 7 and 8 the freeway multi-clearance distribution delineated for low-density states coincides with the distribution (35), which confirms the surmise that vehicles

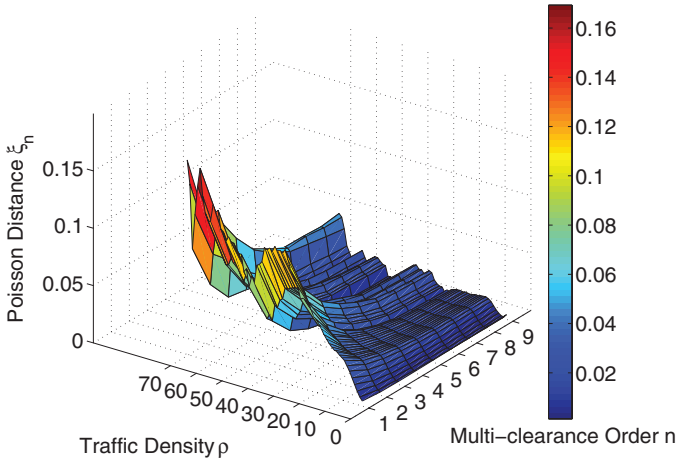

Figure 9. Statistical distance (46) between the empirical multi-clearance distribution and the Poissonian distribution (35). The values ξ_n demonstrate how empirical data differ from ideal independent multi-clearances.

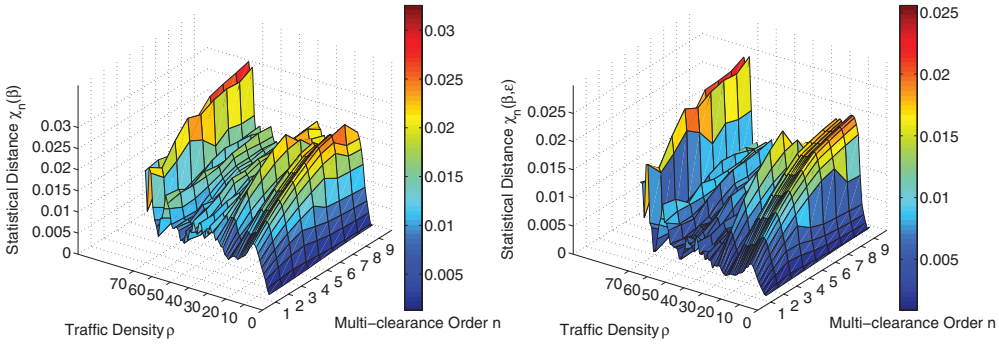

Figure 10. Statistical distance (46) between theoretical and empirical distributions. In the left sub-plot, we display the values of statistical distances $\chi_n(\beta_{\text{fit}})$ between the main-lane multi-clearance distribution $p(t)$ and the zero approximation (38). Analogously, in the right sub-plot we display the values of statistical distances $\chi_n(\beta_{\text{fit}}, \varepsilon_{\text{fit}})$ between the main-lane multi-clearance distribution $p(t)$ and the final approximating formula (47).

in a free-traffic regime are moving as independent elements. Thus, their statistics are purely Poissonian. As the traffic density rises, one can detect (see figure 9) larger deviations from (35), which demonstrates stronger interactions among the cars. Roughly speaking, the fitted thermal parameter β is increasing (in both cases: zero approximation and also the final probability distribution) with the traffic density (see figures 7 and 8). The regions of the temporal descent in the β value agree with the critical regions in the fundamental diagram. To be precise, the transmission between traffic regimes (from a free to a congested regime) causes the transient consolidation of traffic, which leads to a temporal reduction of the driver's mental strain. Since (see [26]) the thermal parameter reflects a level of mental pressure, the detected drop in the course of $\beta = \beta(\varrho)$ is expectable. We add that the density-evolution of thermal parameter β corresponds to the behavior of the associated quantity investigated within the scope of [8, 13, 26, 32]. Moreover, in all traffic states the distribution (47) fits the real-road data more impressively than the original approximation (38), as comprehensible. Such a fact is clearly visible in

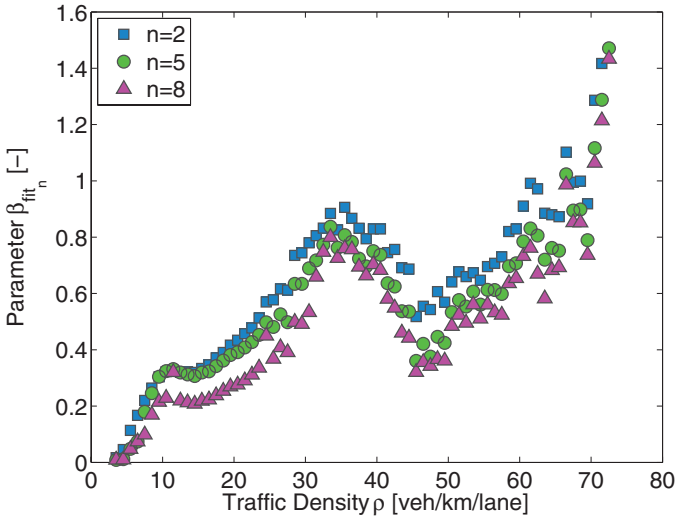

Figure 11. Optimal values of the parameter β as a function of the traffic density. Signs correspond to the values of the parameter β_{fit} fitted according to the relation (48) for the tabulated order of multi-clearance. The perceptible decrease in $\beta_{\text{fit}} = \beta_{\text{fit}}(n)$ reflects the non-zero level of correlations among succeeding clearances, which has not been included in the presented analytical calculations.

figure 10 where the statistical distance $\chi(\beta, \varepsilon)$ between theoretical and empirical distributions are outlined. We add that the larger deviations detected for high-density traffic are caused predominantly by a lack of related data, which reflects in non-smooth empirical histograms.

The important component of our research is a discussion about the stability of the above-introduced fitted procedure, i.e. the stability of the detected interaction-parameter β. The explorations of the dependence between the estimated interaction-parameter β_{fit} and order n of cumulated clearances show (as plotted in figure 11) a slight trend in dependence $\beta_{\text{fit}} = \beta_{\text{fit}}(n)$. Detected descent is a logical consequence of non-zero correlations among succeeding time gaps (not taken into account in the theoretical derivations) and also reflect the fact that outlying cars (with many go-between-vehicles) behave independently (see also figure 9). In spite of this, the estimated interaction-parameter β_{fit} represents an alternative traffic quantity whose values reflect an actual traffic state.

Again, we emphasize that all the statistical tests discussed in this paper have been executed for the main-lane traffic data only. As demonstrated in figures 5 and 6 the time multi-clearance distributions detected in the fast-lane data show the well-known discrepancies with the main-lane data (predominantly in the regions of low densities). For denser streams, the fast-lane distributions converge to the main-lane distributions, which endorses the belief of scientists that in congested traffic states, the correlations among the vehicles in different traffic lanes are much stronger than in free-flow regimes. As is well known, the observed increase of correlation between lanes is reflected by the word 'synchronized' in the name 'synchronized flow'.

9. Summary, conclusion and future prospects

To conclude, this paper deals with cumulative time clearances among several subsequent vehicles passing a given point of an expressway. More specifically, we concentrate predominantly on the time intervals between two following occurrences: (1) the rear bumper

of the kth car has intersected the detector line, (2) the front bumper of the $(k+n+1)$th car has intersected the detector line. Furthermore, we subtract the time periods when vehicles occupy the detectors, which means that a fundamental quantity for this research is the time multi-clearance. Using the local thermodynamical traffic model (originally introduced in [8] and solved analytically in [13]) with short-range repulsive potential among the elements, we have derived (applying the theory of functional convolutions) convenient candidates for the time multi-clearance distribution. A noticeable profit of such calculations can be seen in the fact that the final formula meets all theoretical criteria established by means of the mathematical characteristics of the system observed. This final theoretical distribution represents a one-parametric family of functions, where the one and only fitting parameter β reflects the relevant flow and density. The obtained analytical predictions have been successfully compared with the statistics of freeway data. The detected correspondence between theoretical and empirical distributions allows an elaborated insight into changes of vehicular-traffic microstructure influenced by the momentary traffic state.

In spite of the existing conviction that time-headway distributions in different phases are markedly different (e.g. [5, 9–11]), the stochastic analysis of multi-clearances reveals a different view: the statistical distribution of traffic clearances (and multi-clearances) in all traffic phases belongs to the same one-parametric family of distributions. Those distributions vary from the Poisson distribution (detected for low-density states) to the low-variance distribution (47) that ascertains a presence of stronger correlations among neighboring vehicles, i.e. a stronger vehicular synchronization than in free-flow states. The existing conviction on two types of traffic headway distributions is probably a consequence of insufficient sorting of freeway data. In the earlier scientific papers (dealing with traffic headways), authors usually separate vehicular data into two/three parts only (according a traffic phase) and in fact they combine all the congested states into a single ensemble. But, as is apparent from this research, the congested-phase-distributions significantly vary according to the traffic density. For that reason, the binary data-separation seems to be deficient.

Another substantial outcome of these considerations is the fact that the interaction-parameter β (quantifying a measure of vehicular synchronization) represents an alternative traffic quantity, which reflects an actual traffic state. Moreover, the value β is (in view of the fact that all clearances are scaled) directly connected to the statistical variance of traffic clearances; its estimated value can be therefore obtained by a simple comparison of the empirical and theoretical variances, which is quite effortless.

However, the open problem remains how to approximate the multi-clearance distribution for free-flow vehicles moving in fast lanes. The presence of a large amount of vehicular *leaders* in free-traffic regimes, i.e. a comparable percentage of leaders and *followers*, causes the clearance distribution to probably be a compound of two partial distributions (the first one for leaders, the second one for followers). Thus, the expected approach leading to analytical predictions for fast-lane clearances can be found in the theory of finite mixture distributions (similar to the semi-Poissonian model discussed in [5]).

Acknowledgments

The author would like to thank Cecile Appert-Rolland (Laboratoire de Physique Théorique Université de Paris-Sud, Orsay) for valuable remarks which have been conducive to the presented research. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the project MSM 6840770039 and by the Czech Technical University within the project SGS12/197/OHK4/3T/14.

References

[1] Chowdhury D, Santen L and Schadschneider A 2000 *Phys. Rep.* **329** 199
[2] Helbing D 2001 *Rev. Mod. Phys.* **73** 1067
[3] Hoogendoorn S and Bovy P 2001 *Proc. Inst. Mech. Eng.* I **215** 283
[4] Kerner B S 2004 *The Physics of Traffic* (Berlin: Springer)
[5] Buckley D J 1968 *Transp. Sci.* **2** 107
[6] Krbálek M, Šeba P and Wagner P 2001 *Phys. Rev.* E **64** 066119
[7] Knospe W, Santen L, Schadschneider A and Schreckenberg M 2002 *Phys. Rev.* E **65** 056133
[8] Krbálek M and Helbing D 2004 *Physica* A **333** 370
[9] May A D 1990 *Traffic Flow Fundamentals* (Englewood Cliffs, NJ: Prentice-Hall)
[10] Luttinen R T 1992 *Statistical Properties of Vehicle Time Headways (Transportation Research Record)* (Washington, DC: Transportation Research Board)
[11] Helbing D, Treiber M and Kesting A 2006 *Physica* A **363** 62
[12] Treiber M, Kesting A and Helbing D 2006 *Phys. Rev.* E **74** 016123
[13] Krbálek M 2007 *J. Phys. A: Math. Theor.* **40** 5813
[14] Abul-Magd A Y 2007 *Phys. Rev.* E **76** 057101
[15] Šurda A 2008 *J. Stat. Mech.* P04017
[16] Treiber M and Helbing D 2009 *Eur. Phys. J.* B **68** 607
[17] Li L, Wang F, Jiang R, Hu J and Ji Y 2010 *Chin. Phys.* B **19** 020513
[18] Chen X, Li L, Jiang R and Yang X 2010 *Chin. Phys. Lett.* **27** 074501
[19] Krbálek M and Hrabák P 2011 *J. Phys. A: Math. Theor.* **44** 175203
[20] Reshef D N, Reshef Y A, Finucane H K, Grossman S R, McVean G, Turnbaugh P J, Lander E S, Mitzenmacher M and Sabeti P C 2011 *Science* **334** 1518
[21] Simon N and Tibshirani R 2011 Comment on 'Detecting novel associations in large data sets', unpublished (www-stat.stanford.edu/~tibs/reshef/comment.pdf)
[22] Székely G J and Rizzo M L 2009 *Ann. Appl. Stat.* **3** 1236
[23] Helbing D and Treiber M 2003 *Phys. Rev.* E **68** 067101
[24] Krbálek M 2008 *J. Phys. A: Math. Theor.* **41** 205004
[25] Appert-Rolland C 2009 *Phys. Rev.* E **80** 036102
[26] Krbálek M and Šeba P 2009 *J. Phys. A: Math. Theor.* **42** 345001
[27] Jin X, Zhang Y, Wang F, Li L, Yao D, Su Y and Wei Z 2009 *Transp. Res.* C **17** 318
[28] Tilch B and Helbing D 2000 *Traffic and Granular Flow 99: Social, Traffic, and Granular Dynamics* (Berlin: Springer) p 333
[29] Savitzky A and Golay M J E 1964 *Anal. Chem.* **36** 1627
[30] Mehta M L 2004 *Random Matrices* 3rd edn (New York: Academic)
[31] Sopasakis A 2004 *Physica* A **342** 741
[32] Krbálek M 2010 *Kybernetika* **46** 1108
[33] Huang W and Leimkuhler B 1997 *SIAM J. Sci. Comput.* **18** 239
[34] Kirkpatrick S, Gelatt C D and Vecchi M P 1983 *Science* **220** 671
[35] Krbálek M and Kittanová K 2011 *Procedia Soc. Behav. Sci.* **20** 398
[36] Cline D B H and Samorodnitsky G 1994 *Stoch. Process. Appl.* **49** 75
[37] Teugels J L 1975 *Ann. Probab.* **3** 1000