TWO METHODS FOR ESTIMATING NUMBER OF INTERACTING VEHICLES

SZABOVÁ Zuzana (CZ), KRBÁLEK Milan (CZ)

Abstract. Knowledge of an interaction range in particle systems, especially in vehicular traffic could significantly contribute to modeling of traffic flow . Combination of simulation methods, analytical predictions of headway distribution, and correlation analysis led to several remarkable observations. We observe, that interaction range depends on both resistivity and type of repulsive potential. Moreover we introduce a novel method for detection of number of actively followed vehicles based on perturbation function.

Keywords: Headway distribution, interaction in vehicular traffic, correlation analysis

Mathematics subject classification: Primary 62H20; Secondary 62G07, 62P99

1 Introduction

Increasing traffic demand requires finding either numerical or theoretical models, capable to predict traffic scenarios such as traffic congestions. Recent studies and observations (e.g. [3], [5], [7], [8]) suggest that generally used premise on a short range of traffic interactions do not correspond to traffic reality. However, most traffic models [1], discrete as well as continuous presume that behavior and decision making of a driver is influenced by neighboring vehicles only. By means of correlation analysis we try to determine interaction range with restriction to fast lane and congested traffic phase, with intention to avoid overtaking and transitions between free and congested traffic phases. Moreover we aim to show that interaction range depends on quantities like traffic flow, traffic density or velocity and that correlation between traffic density and velocity plays a role in vehicular data. Further we suggest novel approach for detection of number of actively followed vehicles based on perturbation convolution function.

2 Data-sets and Processing

The used data-sets were provided by the Road and Motorway Directorate of the Czech Republic at the Expressway R1 in Prague, the Czech Republic. Data records contain single vehicle data recorded by double-loop detectors positioned under the expressway R1. For purposes of this research only fast-lane data have been taken into account. Following reasons lead us to this restriction. Firstly, very low

proportion of long vehicles such as lorries or trucks and slow cars is present and more importantly, vehicles in fast lane cannot be overtaken, which could significantly influence traffic flow, especially in a transition phase between free flow and congested flow, when synchronization between lanes may occur. Measured microscopic quantities have been suited into a set of the form

$$S_j = \left\{ (\tau_k^{(in)}, \tau_k^{(out)}, v_k) \in T^{(in)} \times T^{(out)} \times V \mid k = (j-1)m+1, (j-1)m+2, \ldots, jm \right\}, \quad (1)$$

where $j \in M = \left\lfloor \frac{N}{m} \right\rfloor$, $m = 50$ is the fixed sampling size and $N = \# \left\{ \tau_k^{(in)} \mid k \right\}$. Sets $T^{(in)}$ and $T^{(out)}$ are defined as follows.

$$T^{(in)} = \left\{ \tau_k^{(in)} \in \mathbb{R}_0^+ \mid k = 1, 2, \ldots, N \right\}$$

includes the chronologically-ordered times when the front bumper of k-th car has intersected the detector line. Respectively

$$T^{(out)} = \left\{ \tau_k^{(out)} \in \mathbb{R}_0^+ \mid k = 1, 2, \ldots, N \right\}$$

includes the chronologically-ordered times when the rear bumper of k-th car has left the detector. Velocities of succeeding vehicles have been suited into a set

$$V = \left\{ v_k \in \mathbb{R}_0^+ \mid k = 1, 2, \ldots, N \right\},$$

where v_k is a velocity of the k-th vehicle. For each sub-sample S_j, where $j = 1, 2, \ldots, m$, macroscopic quantity such as local flux (intensity)

$$Q_j = \frac{m}{\tau_{jm}^{(out)} - \tau_{(j-1)m+1}^{(in)}}, \quad (2)$$

representing number of vehicles passing the detector per fixed time interval (most commonly 1 hour) or average velocity

$$\bar{v}_j = \frac{1}{m} \sum_{k=(j-1)m+1}^{jm} v_k \quad (3)$$

can be calculated. These quantities can be further used to determine another macroscopic quantity, local vehicular density

$$\rho_j = \frac{Q_j}{\bar{v}_j}, \quad (4)$$

Which can be presented as number of vehicles passing the detector per fixed space unit (most commonly 1 km). Subsequently, time-clearances are scaled to sample-means equal to one and they are defined by[5]

$$t_k = \frac{m \left(\tau_k^{(in)} - \tau_k^{(out)} \right)}{\sum_{i=(\lceil k/m \rceil - 1)m+1}^{\lceil k/m \rceil m} \tau_i^{(in)} - \sum_{i=(\lceil k/m \rceil - 1)m+1}^{\lceil k/m \rceil m} \tau_i^{(out)}}. \quad (5)$$

We will profit from this re-scaling procedure when comparing empirical headway distributions to theoretical predictions having an estimated value equal to one.

At the very beggining we assume that in regions of congested traffic the interactions are more intense than in a free flow phase (which is meaningful premise), in which drivers are driving at their desired speed with desired time clearances. Earlier studies [10] show that both velocities correlations and time clearance correlations depend on traffic flow as well as on the traffic density. Moreover we would like to eliminate mixing of different traffic phases. Therefore we perform segmentation by local density ρ and flux Q (according to [4], [5]).

3 Segmentation of vehicular data by density and velocity

Study of the interaction range will be carried out separately for each flux-density region/window summarized in table 1 as in [11].

w	$\rho_{\min}$ [veh.km^{-1}]	$\rho_{\max}$ [veh.km^{-1}]	$Q_{\min}$ [veh.h^{-1}]	$Q_{\max}$ [veh.h^{-1}]	$\beta_w^{\log}$	β_w^{hyp}
1	25	30	500	2000	1.50	0.35
2	30	35	500	2500	2.55	0.95
3	35	40	500	3000	2.65	1.00
4	40	45	500	3000	3.15	1.30
5	45	50	500	3000	3.40	1.45
6	50	55	500	2500	3.65	1.60
7	55	60	500	2500	3.60	1.60
8	60	65	500	2500	4.50	1.90

Tab. 1. Basic information on the vehicular data used.

We also perform segmentation by velocity regions. Study of the interaction range will be carried out separately for each velocity region summarized in Tab. 2.

w	$\bar{v}_{j,\min}$ [km.h^{-1}]	$\bar{v}_{j,\max}$ [km.h^{-1}]	$\beta_w^{\log}$	β_w^{hyp}
1	20	30	2.90	1.45
2	30	35	3.50	1.35
3	35	40	3.75	1.50
4	40	45	3.25	1.35
5	45	50	3.00	1.35
6	50	55	3.80	1.25
7	55	60	2.90	1.15
8	60	70	2.45	0.95

Tab. 2. Basic information on the vehicular data used.

Basic insight into vehicular interactions can be obtained by means of correlation analysis. Two approaches have been considered. Pearson's correlation coefficient however disposes disadvantage that can obtain even negative values implying anti-correlation and more importantly zero value does not imply statistical independence. Brownian correlation coefficient [9] does not suffer from these shortcomings as has been proven in [9].

Definition 1 *Let X, Y denote random vectors with finite second moments. Then Brownian distance*

correlation $\mathscr{R}_{\mathscr{B}}(X,Y)$ is defined by

$$\mathscr{R}^2_{\mathscr{B}}(X,Y) \begin{cases} \dfrac{\mathscr{V}^2_n(X,Y)}{\sqrt{\mathscr{V}^2_n(X)\mathscr{V}^2_n(Y)}}, & \mathscr{V}^2_n(X)\mathscr{V}^2_n(Y) > 0; \\ 0, & \mathscr{V}^2_n(X)\mathscr{V}^2_n(Y) = 0, \end{cases} \qquad (6)$$

where $\mathscr{V}_n(X,Y)$ denotes distance covariance and $\mathscr{V}_n(X), \mathscr{V}_n(Y)$ represent the associated variances.

However, this type of correlation coefficient also comes with deficiencies, such as strongly correlated variables are not causatively linked. Since, we will not be dealing with strongly correlated variables as was observed in [10], we will dismiss this disadvantage. From the definition of the brownian correlation coefficient it is clear that we will be working with vectors of random variables. Therefore we define

$$X_w := \{x^w_1, x^w_2, \ldots, x^w_\ell\} \qquad (7)$$

and

$$X^{(n)}_w := \{x^w_{1+n}, x^w_{2+n}, \ldots, x^w_{\ell+n} \mid n \in \mathbb{N}\}, \qquad (8)$$

Which are the sets of chronologically ordered individual clearances associated to the w-th flux-density window. Now we can proceed to evaluation of Brownian correlation coefficient $\mathscr{R}_{\mathscr{B}}(X_w, X^{(n)}_w)$ with respect to number of separating vehicles. It follows from Fig. 1 that values of distance-correlation coefficient are mostly decreasing with n in all density regions as well as in all velocity regions.
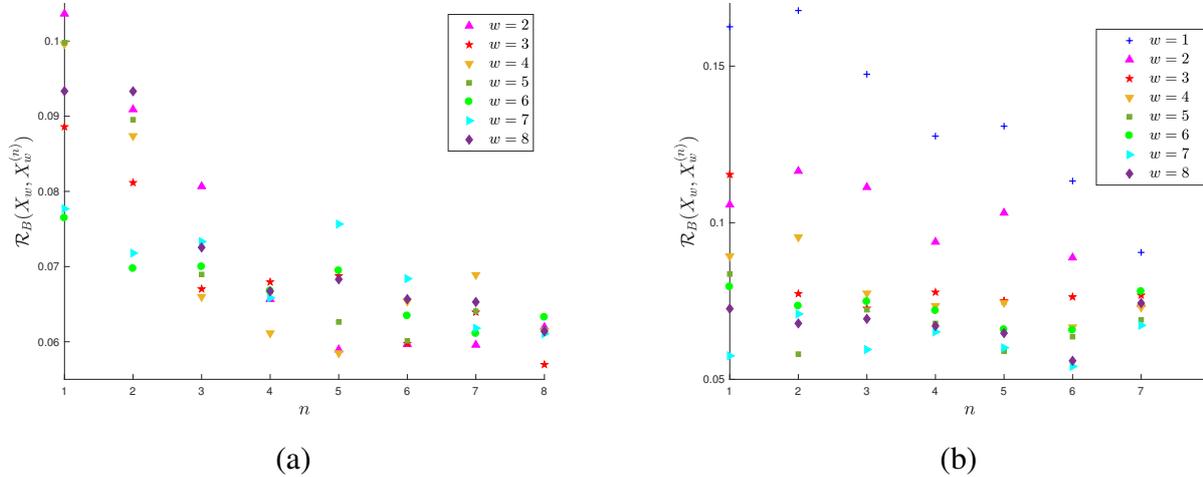

(a)　　　　　　　　　　　　(b)

Fig. 1. rownian distance-correlation of clearances in different density regions (see figure (a)) and velocity regions (see figure (b)).

Strongest correlation is detected for $n = 1$, means that clearances between neighboring vehicles are more strongly correlated compared to clearances measured between more distant vehicles (as in [3]). This is true especially in case of segmentation by density. Moreover, the lower the velocity, the hire the correlations are. As expected, when one finds himself in a strongly congested traffic, drivers are driving bumper to bumper, which generates strong correlation between clearances. On the contrary, when driving in a free phase, drivers tend to maintain desired gap to their predecessors, which

corresponds to very little synchronization and therefore very little correlation between clearances. The isolated application of distance-correlations of clearances is as understandable ineffective for purposes of interaction range estimation. Therefore we focus our attention to the generalized local thermodynamic model allowing middle-ranged interactions. Main characteristics of the model are statistical resistivity β and thermodynamic potential. Short-ranged version of this model applying hyperbolic potential has been solved in [4]. One of the main results of the article is determination of the respective headway distribution, which reads

$$p(x) = \Theta(x) A e^{-\frac{\beta}{x}} e^{-Dx}, \tag{9}$$

where

$$D = \beta + \frac{3 - \sqrt{\beta}}{2}, \tag{10}$$

$$A = \left(2\sqrt{\frac{\beta}{D}} \mathscr{K}_1(2\sqrt{\beta D}) \right)^{-1}. \tag{11}$$

Here $\Theta(x)$ stands for Heaviside unit-step function and $\mathscr{K}_\nu(x)$ for Macdonald's function of the ν-th order. Analogously in article [4] has been determined a headway distribution for the thermodynamic model with short-ranged logarithmic potential, which reads

$$\eta(x) = \Theta(x) \frac{(\beta+1)^{\beta+1}}{\Gamma(\beta+1)} x^\beta e^{-(\beta+1)x}. \tag{12}$$

Empirical values of resistivity β summarized in table 1 (columns 6 and 7) have been determined by means of MDE (minimum distance estimator)

$$\beta_w^h := \mathrm{argmin}_{\beta \in [0,\infty)} \int_0^\infty \left| H_w(x) - \Theta(x) A(\beta) e^{-\frac{\beta}{x}} e^{-D(\beta)x} \right| \mathrm{d}x,$$

$$\beta_w^{log} := \mathrm{argmin}_{\beta \in [0,\infty)} \int_0^\infty \left| H_w(x) - \Theta(x) \frac{(\beta+1)^{\beta+1}}{\Gamma(\beta+1)} x^\beta e^{-(\beta+1)x} \right| \mathrm{d}x,$$

minimizing the statistical distance between theoretical prediction (9), (12) and histogram function $H_w(x)$, respectively. Steady states of the model with middle-ranged hyperbolic potential are not analytically calculated at the moment for any value of resistivity β. In contrast, steady states have been determined for model with middle-ranged logarithmic potential and specific values of resistivity β in [2] and [12]. However, numerical representation allows obtaining all the steady-distributions. Output of the model is vector of steady-state headways/clearances/distances (these terms are equivalent because particles are dimensionless) with average value equal to one

$$X = (x_1, x_2, \ldots, x_N),$$

where number of particles is equal to length of the circle $N = L$. Simultaneously, we define vector

$$X_n = (x_{n+1}, x_{n+2}, \ldots, x_{n+N}),$$

where n denotes number of particles lying between considered particles. For fixed interaction range and stochastic resistivity one can calculate Brownian distance-correlation coefficient $\mathscr{R}_{\mathscr{B}}^n(X, X_n)$. In Fig. 2 values of correlation coefficients $\mathscr{R}_{\mathscr{B}}^n(X, X_1)$ are plotted with respect to interaction range $k \in$

$\{1, 2, 3, 4, 5, 6\}$ and resistivity β for both hyperbolic and logarithmic potential, respectively. Beside case $k = 1$ yielding statistical independence among neighboring headways, one can detect values of distance-correlation coefficient increasing with the statistical resistivity. To estimate an interaction range, we plot empirical values of statistical resistivity β_w^h, β_w^{log} into these diagrams respectively and search for the closest correlation curve. To eliminate fluctuations, we use polynomial interpolation of correlation curves.

(a)

(b)
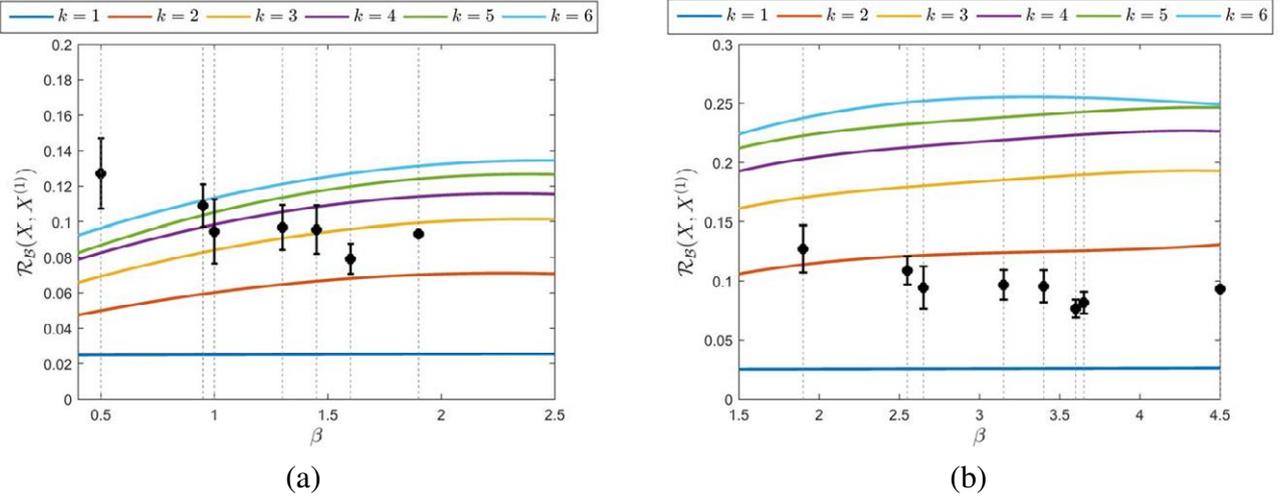
Fig. 2. Curves represent polynomial interpolations of distance-correlation coefficient. Black bullets represent empirical values of Brownian correlation coefficient obtained by MDE for hyperbolic potential (a) and logarithmic potential (b).

It is apparent that interaction range for both logarithmic and hyperbolic potential depends on traffic density. We can summarize that with increasing traffic density (in this case, with increasing statistical resistivity β) interaction range decreases, since correlation drops. In both cases, however, interaction is a middle-ranged in all density regions. Interaction is stronger when using hyperbolic potential. Eventhough we performed several restrictions (only fast lane data from congested phase were used omitting transition phases), we still can not guarantee homogeneity of velocities in each flux-density window. At this point it is convenient to study interactions with respect to velocities.

We use the same approach as for segmentation by velocities. Empirical values of resistivity β (summarized in Tab. 2) have been determined by means of MDE. Similarly to previous section we plot interpolated correlation curves and empirical values of stochastic resistivity in one diagram for both hyperbolic and logarithmic potential (see Fig. 3). In both cases with increasing resistivity, distance-correlation coefficients are increasing. It is also visible that all interactions are middle-ranged, as in the previous section.

4 Perturbation function

At this point it is essential to define multi-clearances of the order $\mu \in \mathbb{N}$ as follows

$$x_k \mid \mu = x_k + x_{k-1} + \cdots + x_{k-\mu}. \tag{13}$$

If the interactions in the system are short-ranged only, then random variables $x \mid \mu_1$ and $x \mid \mu_2$, where $\mu_1 \neq \mu_2$ are independent and the distribution of clearances among $\mu + 2$ particles is given by
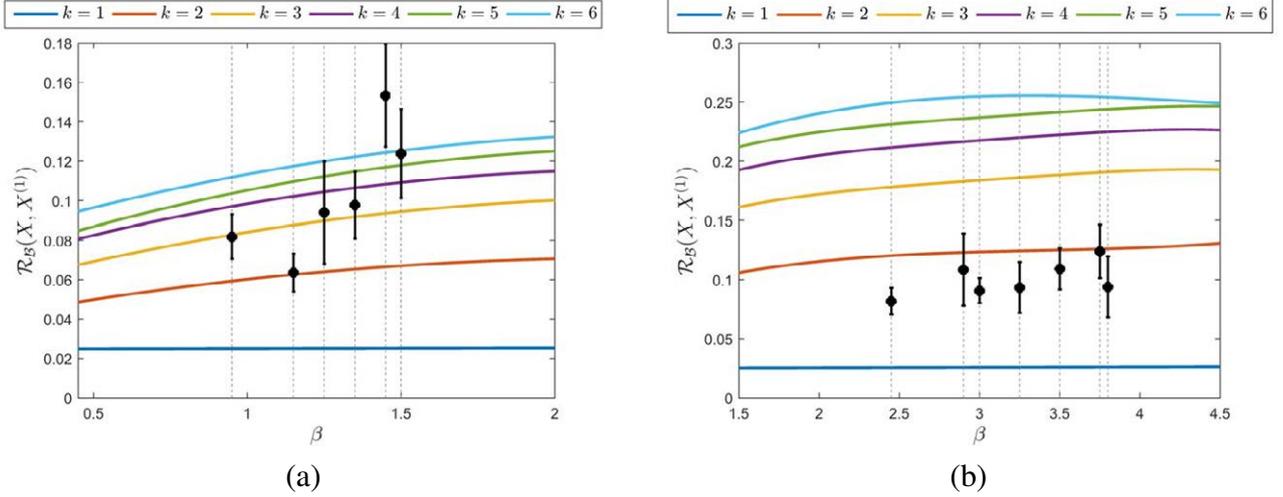
Fig. 3. Curves represent polynomial interpolations of distance-correlation coefficient. Black bullets represent empirical values of Brownian correlation coefficient obtained by MDE for hyperbolic potential (a) and logarithmic potential (b).

convolution formula

$$\rho(x \mid \mu) = \rho(x) \star \rho(x \mid \mu - 1) \equiv \int_{\mathbb{R}} \rho(s)\rho(x - s \mid \mu - 1)\mathrm{d}s. \tag{14}$$

Unfortunately, this formula holds true only for ensembles with short-ranged forces only. Therefore we introduce two-parametric family of probability densities defined in [6]

$$p(x) = \Theta(x)Ax^{\alpha}\mathrm{e}^{-\frac{\beta}{x}}\mathrm{e}^{-Dx}, \tag{15}$$

where

$$D = \frac{\alpha}{\mu} + \frac{\beta}{\mu^2} + \frac{3 - \mathrm{e}^{-\sqrt{\frac{\beta}{\mu}}}}{2\mu}, \tag{16}$$

$$A = \left(2\sqrt{\frac{\beta}{D}}\mathscr{K}_{\alpha+1}(2\sqrt{\beta D})\right)^{-1}. \tag{17}$$

Here $\Theta(x)$ stands for Heaviside unit-step function and $\mathscr{K}_{\nu}(x)$ for Macdonald's function of the ν-th order. To estimate the parameter α, β we use MDE method (Minimum Distance Estimator) given by formula

$$(\hat{\alpha}, \hat{\beta}) := \mathrm{argmin}_{(\alpha,\beta)\in C_{\sigma}} \left(\int_{\mathbb{R}} \mid H(x \mid \mu) - \hat{\rho}(x \mid \mu) \mid^2 \mathrm{d}x\right)^{1/2} \tag{18}$$

$$C_{\sigma} = \left\{(\alpha, \beta) \in \Omega : \int_0^{\infty} x^2 \rho(x \mid (\alpha, \beta))\mathrm{d}x = \mu^2 + \sigma^2\right\}, \tag{19}$$

where $H(x \mid \mu)$ is histogram function (distribution of clearances given by Metropolis-Hastings algorithm) and $\hat{\rho}(x \mid \mu)$ is MDE-estimation. In Fig. 4 we plot estimated values of parameters $\hat{\alpha}, \hat{\beta}$ with respect to degree of multi-clearance. We conclude, that with increasing degree of multi-clearance one obtains increasing values of parameters $\hat{\alpha}, \hat{\beta}$.
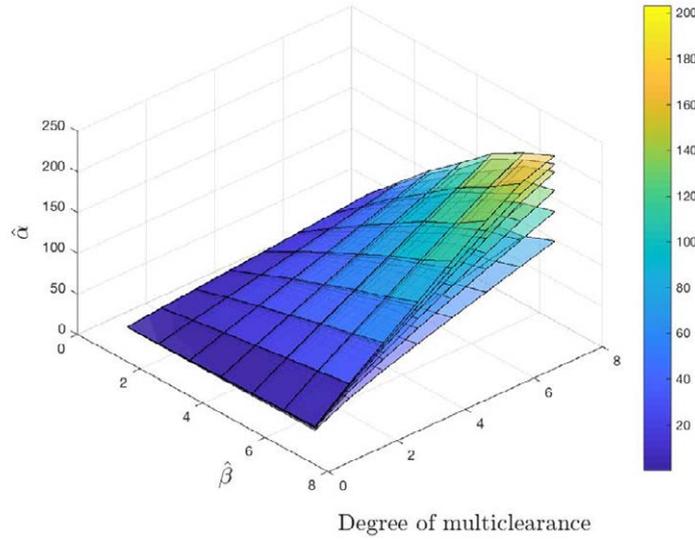
Fig. 4. Estimated values of parameters $\hat{\alpha}, \hat{\beta}$ with respect to degree of multi-clearance for different interactions, starting with short-ranged interaction (lowest plane) and increasing up to interaction of 7 vehicles (upper plane).

Further we would like to introduce a method for detection of interaction range. As we mentioned before, we can use a convolution rule for estimation of probability density $\rho(x \mid \mu), (\mu \neq 1)$ if random variables x and $x \mid \mu - 1$ are statisticaly independent. Therefore we introduce perturbation function

$$\psi(x \mid \mu) = \int_0^x (\hat{\rho}(y \mid \mu) - \hat{\rho}(y \mid 1) \star \hat{\rho}(y \mid \mu - 1)) \mathrm{d}y, \tag{20}$$

which is used for testing statistical independence of different multi-clearances. Courses of perturbation functions for different values of stochastic resistivity β, interaction range and degree of multi-clearance are shown in Fig. 5.

It is desired to study how the perturbation function $\psi(x \mid \mu)$ deviates from zero value. For intentions of this quantification we use Kolmogorov distance defined by formula

$$G(\mu) = \sup_{x \in \mathbb{R}} \mid \psi(x \mid \mu) \mid . \tag{21}$$

Course of Kolmogorov distance with respect to interaction range, resistivity β and degree of multi clearance is shown on Fig. 6. Values of Kolmogorov distance are low for all interaction ranges and the associated neighboring planes are very close to each other. Kolmogorov distance planes have been created by means of simulation method (Metropolis-Hastings algorithm). Now question arises how precisely we are able to determine course of this planes. In Fig. 7(a) it is shown Kolmogorov distance plane with error bars for model with middle-ranged interaction (interaction of 5 vehicles) and in Fig. 7 (b) it is shown Kolmogorov distance plane with error bars for model with middle-ranged interaction (interaction of 5,6,7 vehicles, respectively). Since the upper error bar is intersecting Kolmogorov distance plane for interaction of 7 vehicles, at this point we would not be able to determine a range of interaction. Unambigously, this is true only for planes with interaction of more than 6 vehicles. For lower interaction the error bars are not intersecting neighboring distance planes.
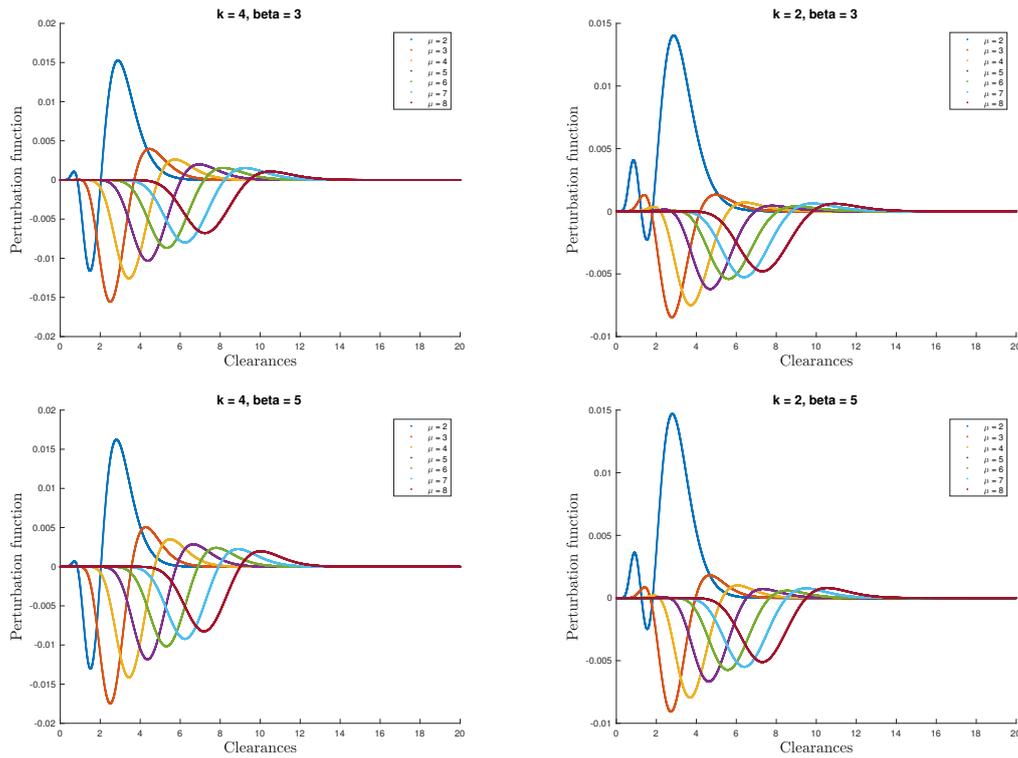
Fig. 5. Courses of perturbation functions for different values of stochastic resistivity β, interaction range and degree of multi-clearance.
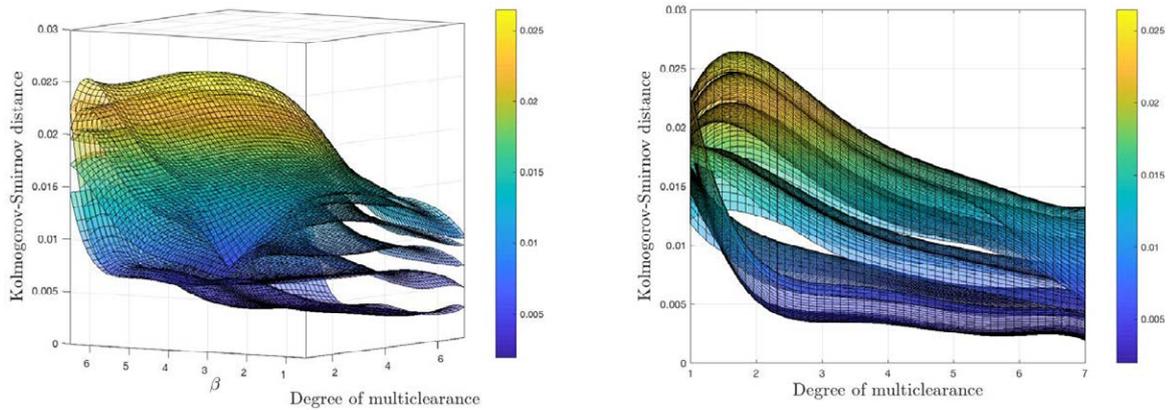

Fig. 6. Courses of Kolmogorov distance of perturbation function with respect to stochastic resistivity β, interaction range and degree of multi-clearance.

5 Conclusion

Main task of this paper was to estimate a number of actively followed vehicles by means of Brownian distance-correlation coefficients and numerical data obtained from simulations of one-dimensional particle gas with both middle-ranged hyperbolic and logarithmic potential. Segmentation of the empirical vehicular data by density and velocity has been performed. These separations lead to contradictory results. When segmentation by local density was used, we observed that with increasing stochastic resistivity, correlations were decreasing. On the contrary, segmentation by velocities lead to increasing correlations with increasing stochastic resistivity. This means, that either one of the seg-
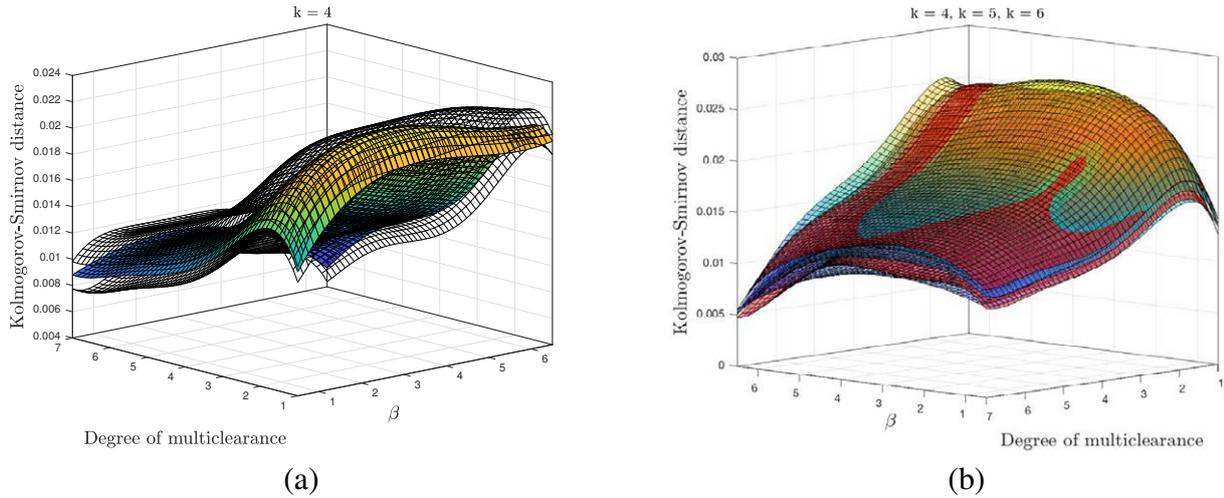
Fig. 7. Course of Kolmogorov distance plane with error bars (transparent planes) for interaction of 5 vehicles (a) and course of Kolmogorov distance plane with error bars (red planes) for interaction of 5,6,7 vehicles respectively from down to up (b).

mentation methods is not appropriate or that a correlation between velocities and densities is present. Despite the contradictory results, we observe that in both cases the interactions are at least middle-ranged.

Method based on perturbation function needs to be tested and validated on empirical data. From simulations it is obvious so far, that for higher interaction range, Kolmogorov distance planes are too close to each other and therefore it will be difficult to determine precise interaction range (if it comes to high order interactions). On the other hand, it should not be difficult to determine at least character of interactions (short ranged, middle ranged), since Kolmogorov distance plane for short-range interaction is far from the rest of the planes derived for middle-ranged interactions (for $\mu > 1$).

References

[1] D. HELBING. Traffic and related self-driven many-particle systems. *Rev. Mod. Phys.*, 73, 1067, 2001.

[2] E. B. BOGOMOLNY, U. GERLAND, C. SCHMIT. Short-range plasma model for intermediate spectral statistics. *Eur. Phys. J. B*, 19, 121, 2001.

[3] L. NEUBERT, L. SANTEN, A. SCHADSCHNEIDER, M. SCHRECKENBERG. Single-vehicle data of highway traffic: A statistical analysis. *Phys. Rev. E*, 60, 6480, 1999.

[4] M. KRBÁLEK. Equilibrium distributions in a thermodynamical traffic gas. *J. Phys. A: Math. Theor.*, 40, 5813, 2007.

[5] M. KRBÁLEK. Theoretical predictions for vehicular headways and their clusters. *J. Phys. A: Math. Theor.*, 46, 445101, 2013.

[6] M. KRBÁLEK, J. APELTAUER, T. APELTAUER, Z.SZABOVÁ. Three methods for estimating a range of vehicular interactions. *Physica A*, 491, 112-126, 2018.

[7] C. APPERT-ROLLAND. Experimental study of short-range interactions in vehicular traffic. *Phys. Rev. E*, 80, 036102, 2009.

[8] M. TREIBER, A.KESTING. *Traffic Flow Dynamics*. Berlin: Springer, 2013.

[9] G. J. SZÉKELY, M. L. RIZZO. Brownian distance covariance. *Annals of Applied Statistics*, 3/4, 1236, 2009.

[10] Z. SZABOVÁ. Correlation analysis in ensembles of empirical vehicular data. *Bachelor thesis*, 2016.

[11] Z. SZABOVÁ. Statistical detection of an interaction range in particle systems. *Proceedings of SPMS 2017*, Dobrichovice 2017.

[12] Z. SZABOVÁ. Statistical detection of an interaction range in particle systems. *Research task*, 2017.

Current address

Szabová Zuzana, Bc.
Department of mathematics
Faculty of Nuclear Sciences and Physical Engineering
Czech Technical University in Prague
Trojanova 13, Prague, Czech Republic
E-mail: szabov.zuzana@gmail.com

Krbálek Milan, doc. Mgr., PhD.
Department of mathematics
Faculty of Nuclear Sciences and Physical Engineering
Czech Technical University in Prague
Trojanova 13, Prague, Czech Republic
E-mail: milan.krbalek@fjfi.cvut.cz