Super-random states in vehicular traffic – detection & explanation

Milan Krbálek[1], **František Šeba**[2], **Michaela Krbálková**[2,3]

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

[2] Faculty of Science University of Hradec Králové, Hradec Králové, Czech Republic

[3] Jan Perner Transport Faculty, University of Pardubice, Pardubice, Czech Republic

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

Abstract. This article deals with specific states of traffic flow on a two-lane freeway, in which statistical fluctuations of microscopic quantities (inter-vehicle gaps) are significantly higher than in systems with absolutely random events (Poisson systems). These anomalous states (super-random) are detected in empirical traffic data, specifically in the fast lane at traffic densities up to 25 vehicles per kilometer. The origin of these states is then explained mathematically (using the theory of balance particle systems and tools of random matrix theory), physically (by means of an one-dimensional particle gas subjected to local perturbations caused by overtaking cars) and empirically (using an analogy with phenomena observed in photon counting experiments). In the article we show that overtaking maneuvers, when vehicles from a slow lane are injected into a fast-lane stream of faster moving vehicles, disrupt a local balance in microstructure of fast-lane stream and cause atypical arrangement of vehicular positions, that is very rare, generally. With help of original numerical model we demonstrate that the anomalous states detected are identical to equilibrium states formed in a stochastic particle gas with a potential containing, in addition to a repulsive component, also an attractive component.

Keywords: Vehicular traffic, Super-Poissonian states, Vehicular headway modelling, Statistical rigidity, Vehicular bunching, Random matrix theory

1. Introduction

During the last twenty years, the statistical approaches applied to driven many-particle systems have represented key methods in the discovery, explanation, and description of specific macroscopic and microscopic phenomena in traffic flow [1, 2, 3, 4, 5, 6]. When investigating the microstructure of traffic flows, it has been shown that distributions of microscopic traffic quantities are highly sensitive to small changes in traffic density or intensity [4, 7]. As the values of these macroscopic variables change, both the average values of the micro-quantities and their variances change significantly [8, 9]. This indicates that the parametric space, within which the parameters of the random variables (commonly analyzed in traffic disciplines) are detected, is quite extensive. Fortunately, it appears that when traffic data is analyzed taking into account this high sensitivity to the values of macroscopic variables (i.e. applying 3s-unification procedure presented below) suitable candidates for theoretical distribution curves can be found in a very compact form. In addition, respective probability distributions reflect all the mathematical properties required by the general theory of one-dimensional particle systems [10, 11]. This situation opens up the possibility to answer advanced questions concerning an inner structure of vehicular ensembles and its temporal evolution, which is the content of this paper. For correctness we add that there are a number of other features of vehicular traffic that can not be described by a simplified approach based on the above-mentioned analogy with particle gas. This is due to the fact that vehicular traffic represents, as is well known, a complex physical system.

With increasing availability of single-vehicle data, the number of scientific challenges associated with the traffic microstructure and the time-space evolution of microscopic quantities has been growing. In this text we will try to decipher the essential characteristics of the spatiotemporal arrangement of traffic samples extracted from traffic measurements carried out on two-lane expressways and put them in the context of traffic theory. For these purposes we will use segmentation-oriented analysis of traffic data, selected parts of Random Matrix Theory [12], mathematical theory of balanced particle systems [11], and appropriate physical analogies to the researched topic.

To be specific, in this article we will try to show that traffic flow violates, under certain circumstances, a set of theoretical limits that should not be exceeded in typical particle systems. We show that in a main lane the traffic system at the microscopic level behaves orderly, while the flow in a fast lane shows unexpected statistical anomalies in a certain density region. They originate in more complex interactions between vehicles, which leads to violation of assumptions in relevant mathematical theorems. From a mathematical point of view, these anomalous states represent the so-called *super-Poisson states* (*super-random,* less formally), which were already detected e.g. in photon counting experiments, in measurements of atomic resonance fluorescence, or in quantum optics [13, 14, 15]. In these experiments, thermal light shows intensity fluctua-

tions as its inherent aspect and as a result of the fact, that there is a statistical tendency for photons to arrive simultaneously at a detector. This implies the so-called *photon bunching* – a phenomenon generally attributed to the fact that photons are Bosonic and known as the Hanbury Brown and Twiss effect [16].

The paper is organized as follows. In the following section, we present empirical traffic data analyzed in this article, the method of their measurement and their mathematical structure. In the third section, we briefly mention the creation of a fundamental traffic map. The fourth section focuses on a detailed description of the unification procedure designed for processing data measured on real roads. In the next part, we present the so-called theory of balance particle systems – a mathematical tool for analytical predictions in the researched area. In the sixth section, we classify systems or states of a given system according to the degree of statistical compressibility. Using the instrument of statistical compressibility we then define sub-Poisson and super-Poisson states. The latter represent, from a certain point of view, anomalous statistical states, because the rate of their fluctuations significantly exceeds the limit reached in Poisson systems with absolutely random events. In the seventh section we detect these super-Poisson states in vehicle-by-vehicle data and in the eighth section we explain the reasons for their occurrence.

2. Empirical traffic samples

Single-vehicle data analyzed in this paper have been measured during 90 days at several supersaturated segments of highway circuit R1 Prague, Czech Republic by technology of induction double-loop detection. Data records containing twelve data-files, each having data from measurements lasting three months approximately, have been provided by The Road and Motorway Directorate of the Czech Republic. The total volume of data is approximately 55 millions. Typical property of these data sets is that they have been measured at the same time-fixed location (a so-called *detector line*). Typical outputs of all similar traffic measurements look like

$$T^{(\mathrm{in})} = \{\tau_k^{(\mathrm{in})} \in \mathbb{R} \mid \tau_{k-1}^{(\mathrm{in})} < \tau_k^{(\mathrm{in})} \ \wedge \ \tau_0^{(\mathrm{in})} := 0\}, \tag{1}$$

$$T^{(\mathrm{out})} = \{\tau_k^{(\mathrm{out})} \in \mathbb{R} \mid \tau_{k-1}^{(\mathrm{in})} < \tau_{k-1}^{(\mathrm{out})} \leqslant \tau_k^{(\mathrm{in})} < \tau_k^{(\mathrm{out})} \ \wedge \ \tau_0^{(\mathrm{out})} := 0\}, \tag{2}$$

$$\Upsilon = \{v_k \in \mathbb{R}_0^+ \mid k = 1, 2, \ldots, \mathtt{N}\}, \tag{3}$$

$$\Lambda = \{\ell_k \in \mathbb{R}^+ \mid k = 1, 2, \ldots, \mathtt{N}\}, \tag{4}$$

where $T^{(\mathrm{in})}$ and $T^{(\mathrm{out})}$ collect instants τ_k when a front/rear bumper of a kth car has intersected a detector line and Υ and Λ collects individual velocities v_k and lengths ℓ_k of cars, respectively.

If loop measurements are accompanied by image processing technology then additional data sets

$$\Xi^{(\mathrm{front})} = \{\xi_k^{(\mathrm{front})} \in \mathbb{R} \mid \xi_k^{(\mathrm{front})} < \xi_{k-1}^{(\mathrm{front})} \ \wedge \ \xi_0^{(\mathrm{front})} := +\infty\},$$

$$\Xi^{(\mathrm{rear})} = \{\xi_k^{(\mathrm{rear})} \in \mathbb{R} \mid \xi_k^{(\mathrm{rear})} < \xi_k^{(\mathrm{front})} \leqslant \xi_{k-1}^{(\mathrm{rear})} < \xi_{k-1}^{(\mathrm{front})} \ \wedge \ \xi_0^{(\mathrm{rear})} := +\infty\},$$

collecting positions of front/rear bumpers at the fixed time, are to disposal. However, such a doubled measurements are very rare, which results in the fact that locations ξ_k (in contrast to instants of time τ_k) belong to indirectly determined traffic quantities. Therefore, we refer instants of time τ_k, velocities v_k, and lengths ℓ_k as *primary quantities,* whereas locations ξ_k are referred to as *secondary quantities.* Based on these quantities one can define empirical values of *time headways/clearances* using respective formulas

$$t_k := \tau_k^{(\mathrm{in})} - \tau_{k-1}^{(\mathrm{in})}, \quad z_k := \tau_k^{(\mathrm{in})} - \tau_{k-1}^{(\mathrm{out})}.$$

Distance headways are then defined by

$$s_k := \xi_{k-1}^{(\mathrm{front})} - \xi_k^{(\mathrm{front})}$$

and distance clearances read

$$r_k := \xi_{k-1}^{(\mathrm{rear})} - \xi_k^{(\mathrm{front})}.$$

Graphically, all these quantities are visualized in figure 1. Note that equalities $s_k = v_k t_k$, $r_k = v_k z_k$ are valid only provided that an individual velocity v_k is constant during time interval $[\tau_{k-1}^{(\mathrm{in})}, \tau_k^{(\mathrm{in})}]$, which is less probable assumption, especially when $\tau_k^{(\mathrm{in})} \gg \tau_{k-1}^{(\mathrm{in})}$. It means that loop measurements without additional processing of video/photography generate time clearances and time headways as primary variables, whereas distance headways/clearances represent secondary variables and their approximative values $s_k \approx v_k t_k$, $r_k \approx v_k z_k$ can be therefore burdened by systematic errors. Moreover, one finds

$$r_k = \xi_{k-1}^{(\mathrm{rear})} - \xi_{k-1}^{(\mathrm{front})} + \xi_{k-1}^{(\mathrm{front})} - \xi_k^{(\mathrm{front})} = s_k - \ell_{k-1}.$$

Mathematically, all the quantities from the previous text represent non-negative continuous random variables and are therefore characterized by a standard statistical description using associated probability densities and distribution functions. Consider now a sequence $(\mathcal{X}_k)_{k=1}^{\mathtt{N}}$ of random variables of the same type (clearances, headways). From a statistical viewpoint the empirical headways x_k represent individual realizations of random variables $\mathcal{X}_k$ and one can model respective distributions by standard statistical routines. Generally accepted premise in Vehicular Headway Modelling says that $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_{\mathtt{N}}$ are identically distributed provided that one analyzes homogeneous flows, where macroscopic quantities *(state variables)* are steady in time and a rate of long vehicles is low. To a certain degree of simplification it used to be sometimes speculated that $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_{\mathtt{N}}$ are independent identically distributed (i.i.d.), which is useful simplifying assumption suitable for theoretical calculations [17, 18, 19]. In fact, correlations among headways are significant [20, 21], which points to a more complex interactions among drivers.

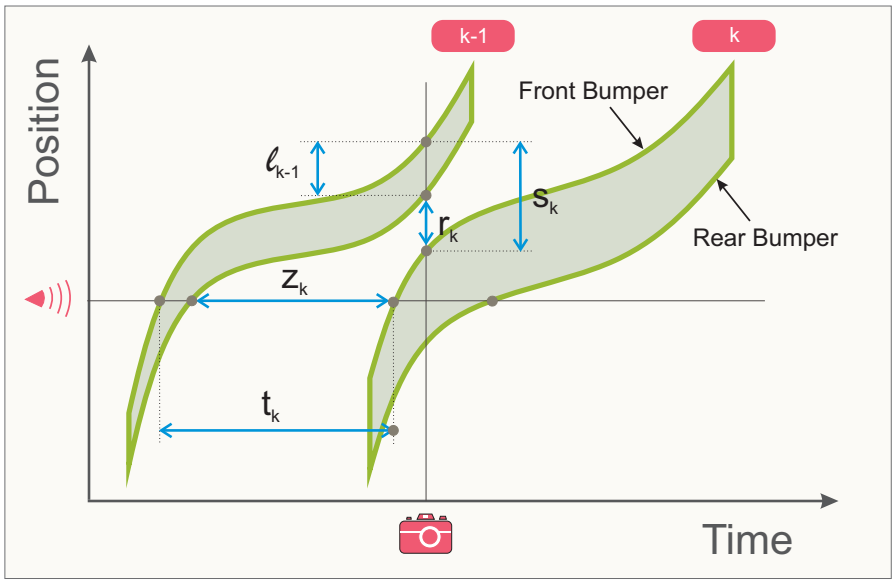

Figure 1. Single-vehicle characteristics obtained by cross-section detectors (marked with z_k and t_k) or by image processing technology (marked with r_k and s_k).

3. Fundamental map of vehicular traffic

As is usual, single-vehicle samples are characterized by the three fundamental traffic quantities: traffic density ϱ, traffic intensity I, or average velocity V. For a given single sample

$$\Xi = \left\{ (\tau_k^{(\mathrm{in})}, \tau_k^{(\mathrm{out})}, v_k, \ell_k) \in T^{(\mathrm{in})} \times T^{(\mathrm{out})} \times \Upsilon \times \Lambda : \ k = 1, 2, \ldots, \mathtt{K} \right\} \qquad (5)$$

of $\mathtt{K}$ succeeding cars one can calculate the single-sample intensity $I = \mathtt{K}/(\tau_{\mathtt{K}}^{(\mathrm{out})} - \tau_1^{(\mathrm{in})})$ and the single-sample mean speed $V = \mathtt{K}^{-1} \sum_{k=1}^{\mathtt{K}} v_k$. Moreover, the hydrodynamics-like expression $\varrho = I/V$ used to be usually accepted as a plausible approximation for the local density [22]. Then, applying the same computational routine for all available data samples $i = 1, 2, \ldots, m$ we acquire binary relations $\Omega_{\mathtt{ID}} = \{(\varrho_i, I_i) : \ i = 1, 2, \ldots, m\}$ and $\Omega_{\mathtt{VD}} = \{(\varrho_i, V_i) : \ i = 1, 2, \ldots, m\}$, whose graphs can be used for synoptic visualizations of macroscopic properties of traffic flow. These graphs represent *fundamental maps* of a given traffic problem.

4. Data processing: 3s-unification procedure

Specific and well-known signs of vehicular traffic are (see [22, 23]): a wide scattering of flow-density data in the congested regime, chaotic evolution of state variables, repetitive sharp increases of density, and a propagation of kinematic waves in a direction opposite to vectors of vehicular velocities. All these effects cause that larger samples of succeeding vehicles show a significant non-homogeneities. Their microstructure is therefore also

non-homogeneous [24], which results in the fact that associated probability distributions are not unimodal and produce, in contrast, courses typical for mixed systems composed from several different distributions.

To gain homogeneous characteristics one has to apply the following 3s-unification procedure that prevents an undesirable mixing of states with different statistical properties (stochastic resistivity and rigidity), different vehicular properties (average headways, clearances, velocities and their variances) and psychological properties (vigilance of drivers, reaction times, decision-making strain).

4.1. Sampling

The first sub-routine in a three-phase unification procedure is *the sampling,* i.e. division of data into homogeneous samples of several neighboring cars.

Consider the data sets (1)–(4), a sampling size $\mathtt{K}$, and number of samples m. Without loss of generality, we assume that $m\mathtt{K} = \mathtt{N}$. For each single sample $i \in \{1, 2, \ldots, m\}$ we denote

$$G_i = \{(i-1)\mathtt{K} + 1, (i-1)\mathtt{K} + 2, \ldots, i\mathtt{K}\}$$

a respective index set and extract relations $\Omega^{(i)}_{\mathtt{ID}}$ and $\Omega^{(i)}_{\mathtt{VD}}$ as introduced in subsection 3. Moreover, we define sample-adjoint sets of individual headways (time and spatial) $T_i := \{t_k : k \in G_i\}$ and $S_i := \{s_k : k \in G_i\}$, sets of individual clearances (time and spatial) $Z_i := \{z_k : k \in G_i\}$ and $R_i := \{r_k : k \in G_i\}$, and sets of velocities $\Upsilon_i := \{v_k : k \in G_i\}$ and lengths $\Lambda_i := \{\ell_k : k \in G_i\}$. From this approach it follows that the ith sample is described by the sample-adjoint values ϱ_i, I_i, V_i and by the random sets $T_i, S_i, Z_i, R_i, \Upsilon_i, \Lambda_i$.

4.2. Scaling

As understandable, the mean values $\langle T_i\rangle, \langle S_i\rangle, \ldots, \langle \Lambda_i\rangle$ of the random data sets $T_i, S_i, Z_i, R_i, \Upsilon_i, \Lambda_i$. can be easily enumerated by means of the values ϱ_i, I_i, V_i. Indeed, it holds

$$\langle T_i\rangle = 1/I_i; \quad \langle S_i\rangle = 1/\varrho_i; \quad \langle \Upsilon_i\rangle = V_i \tag{6}$$

and $\langle R_i\rangle = \langle S_i\rangle - \langle \Lambda_i\rangle$. It means that for traffic micro-quantities the statistical characteristics of the first order are hidden in relations $\Omega_{\mathtt{ID}}$ and $\Omega_{\mathtt{VD}}$. Therefore, a scaling of individual micro-quantities represents no loss of information. Above that, the standard unfolding procedure (usually applied in many statistical studies aiming to reveal a non-trivial stochastic universality like in Random Matrix Theory [12, 9]) includes a scaling procedure as its inherent part.

For empirical/experimental data the scaling should be applied as follows. Traffic micro-quantities are converted to associate scaled alternatives. Such a conversion is here demonstrated on the example of the sample-adjoint set $Z_i := \{z_k : k \in G_i\}$ of time

clearances. Associate set $Y_i := \{y_k : k \in G_i\}$ of scaled clearances is calculated using a definition

$$y_k = \frac{z_k \cdot \mathtt{K}}{\sum_{k \in G_i} z_k} = \frac{z_k}{\langle Z_i \rangle} \quad (k \in G_i), \tag{7}$$

which ensures that $\langle Y_i \rangle = 1$. In analogy, we define scaled spatial clearances by

$$x_k = \frac{r_k \cdot \mathtt{K}}{\sum_{k \in G_i} r_k} = \frac{r_k}{\langle R_i \rangle} \quad (k \in G_i). \tag{8}$$

4.3. Segmentation

The final step of the three-phase unification procedure is forced by the fact that most of traffic variables are significantly changing if the state variables vary. It means that random variable characteristics of the first, second, third, and fourth order (average, variance, skewness, and kurtosis, respectively) strongly depend on actual values of traffic macro-quantities. Naturally, a mixing of states with different values of fundamental variables is significantly undesirable. For this reason, one has to analyze data from small sub-area of a fundamental map only. To be specific, denoting Ψ an arbitrary subset of the fundamental map the *segmentation procedure* selects those samples having values of the fundamental variables belonging to Ψ. Therefore, we introduce the *segmented index set*

$$F_\Psi := \{i = 1, 2, \ldots, m : (\varrho_i, I_i) \in \Psi\}. \tag{9}$$

Then, statistical analysis intended is performed separately for scaled micro-quantities

$$Y_\Psi \times X_\Psi \times \Upsilon_\Psi := \{(y_k, x_k, v_k) \in Y_i \times X_i \times \Upsilon_i : i \in F_\Psi \wedge k \in G_i\} \tag{10}$$

extracted from the *segment* Ψ.

5. Balanced particle systems

5.1. Headway distribution in vehicular streams

As is apparent from considerations presented in Appendix 10.1, as well as from numerous traffic data analyzes [6, 7, 8, 9, 11, 18, 21], spatial and temporal clearances between vehicles are statistically distributed according probability density

$$h(x) = A\Theta(x)\mathsf{e}^{-\beta\varphi(x)-\lambda x}, \tag{11}$$

where $\varphi'(x)$ correspond to a repulsive hard-cored force between succeeding vehicles, preventing vehicle collisions. It means that

$$\varphi'(x) \leqslant 0, \quad \lim_{x \to 0^+} \varphi'(x) = -\infty, \quad \lim_{x \to +\infty} \varphi'(x) = 0. \tag{12}$$

Therefore, regardless of the specific choice of potential $\varphi(x)$ the headway distribution (11) belongs to the specific family of probabilistic densities described immediately below.

5.2. The class of balanced functions

The class of balanced functions $\mathscr{B}$ is the space of piecewise continuous functions $f(x)$ on $\mathbb{R}$ with $\mathrm{Dom}(f) = \mathbb{R}$, $\mathrm{Ran}(f) \subset [0, +\infty)$, $\mathrm{supp}(f) \subset [0, +\infty)$, for which there exists positive number $\varkappa$ so that

$$(\forall \alpha > \varkappa) : \lim_{x\to\infty} f(x)\mathsf{e}^{\alpha x} = +\infty \quad \wedge \quad (\forall \alpha < \varkappa) : \lim_{x\to\infty} f(x)\mathsf{e}^{\alpha x} = 0 \quad \text{(balance axiom)}.$$

The number $\varkappa$ is referred to as *balancing index* and denoted by $\mathrm{inb}(f)$. In fact, belonging of headway distribution to the space $\mathscr{B}$ must be met for all one-dimensional particle systems, where interactions are restricted to a few neighboring particles only [11], which definitively corresponds with real-road traffic. In addition, under the conditions (12) applied to p.d.f. (11) the following implication holds

$$\mu_1 = 1 \quad \Longrightarrow \quad \lambda \geqslant 1 \quad \wedge \quad \mu_2 \leqslant 2, \tag{13}$$

where $\mu_k = \int_{\mathbb{R}} x^k h(x)\,\mathtt{d}x$ is $k-$th statistical moment. Moreover, limiting values are $\lambda = 1$, $\mu_2 = 2$ are obtained for Poissonian system, which is either a system with absolutely non-interacting particles, i.e. $\varphi(x) = 0$, or stochastically irresistible system where $\beta = 0$. These variants are described by the exponential headway distribution $h(x) = \Theta(x)\mathsf{e}^{-x}$, for which therefore the statistical variance $\mathtt{VAR}(\mathcal{X}) = \mu_2 - \mu_1^2$ is equal to one.

5.3. Balanced particle systems

Now we introduce the concept of the so-called Balanced Particle Systems as an effective mathematical instrument for description of statistical properties of vehicular microstructure. *Balanced Particle System (BPS)* is understood as a sequence $(\mathcal{X}_k)_{k=0}^{\infty}$ of random multiheadways, i.e. gaps between the reference particle $k = 0$ and the particle indexed by $k + 1$, defined by

$$\mathcal{X}_k := \sum_{i=0}^{k} \mathcal{R}_i,$$

where headways between succeeding particles $\mathcal{R}_0, \mathcal{R}_1, \mathcal{R}_2, \ldots$ are non-negative, identically distributed via the balanced p.d.f. $g(x) \in \mathscr{B}$ referred usually to as *generator of the BPS,* i.e.

$$\mathcal{R}_0, \mathcal{R}_1, \mathcal{R}_2, \ldots \ \backsim \ g(x).$$

Illustratively, it is visualized in the figure 2. The Balanced Particle System is being termed the *Independent Balanced Particle System (iBPS),* if $\mathcal{R}_0, \mathcal{R}_1, \mathcal{R}_2, \ldots$ are independent variables. In this case

$$\mathcal{X}_k \ \backsim \ \bigstar_{i=0}^{k} g(x),$$

where $\star$ symbolizes a convolution. Besides the above-discussed continuous description the particle systems can be described also via a interval frequency $\mathcal{N}_L$, which represents

the number of particles on an interval of length L lying immediately beyond the reference particle. These continuous and discrete descriptions, i.e. headways vs. frequencies, of the same particle system are, as understandable, firmly linked by mathematical relationships [28, 29, 10].

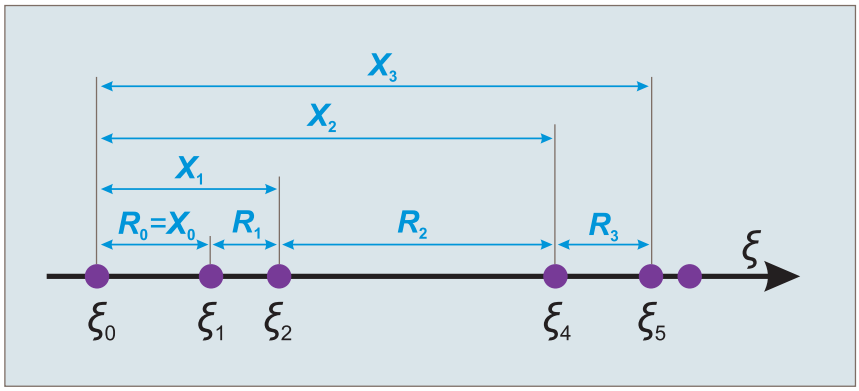

Figure 2. Schematic representation of a headway-sequence $(\mathcal{R}_k)_{k=0}^{\infty}$ and multiheadway-sequence $(\mathcal{X}_k)_{k=0}^{\infty}$.

A privileged place among the mathematical instruments for BPS is occupied by the *statistical rigidity* [12, 25, 26]. This continuous function, frequently used in Random Matrix Theory [12], is defined by

$$\Delta(L) = \sum_{k=0}^{\infty} (k-L)^2 \, \mathbb{P}[\mathcal{N}_L = k].$$

Together with the expected value $\mathtt{E}(\mathcal{N}_L) = \sum_{k=1}^{\infty} k\mathbb{P}[\mathcal{N}_L = k]$ the statistical rigidity represents an extremely useful and robust mathematical tool for classification of the intensity of random fluctuations in one-dimensional systems. It has been successfully used in investigations of vehicular systems as well [3, 6]. For example, previous analyzes of traffic and pedestrian data [27] have quite convincingly distinguished behavior of the rigidity in the both systems, as can be clearly seen in the figure 3. This analysis shows that pedestrian systems are, in contrast to vehicular flows, much more organized. i.e. they are much more closer to deterministic processes than strongly stochastic vehicular transport.

Moreover, it has been shown in [28, 29, 10] that $\Delta(L)$ is linear-like function, for which

$$\Delta(L) \approx \chi L + \delta + o(L^2) \quad (L \to +\infty),$$

where $\chi \geqslant 0$, $\delta \in \mathbb{R}$ are called a *compressibility* [25], and *deflection,* respectively. Analytical studies have proven [10] that the compressibility in iBPC can be expressed by means of the moments of generator $g(x)$ as

$$\chi = \frac{\mu_2 - \mu_1^2}{\mu_1^3} = \frac{\sigma^2}{\mu_1^3}, \tag{14}$$

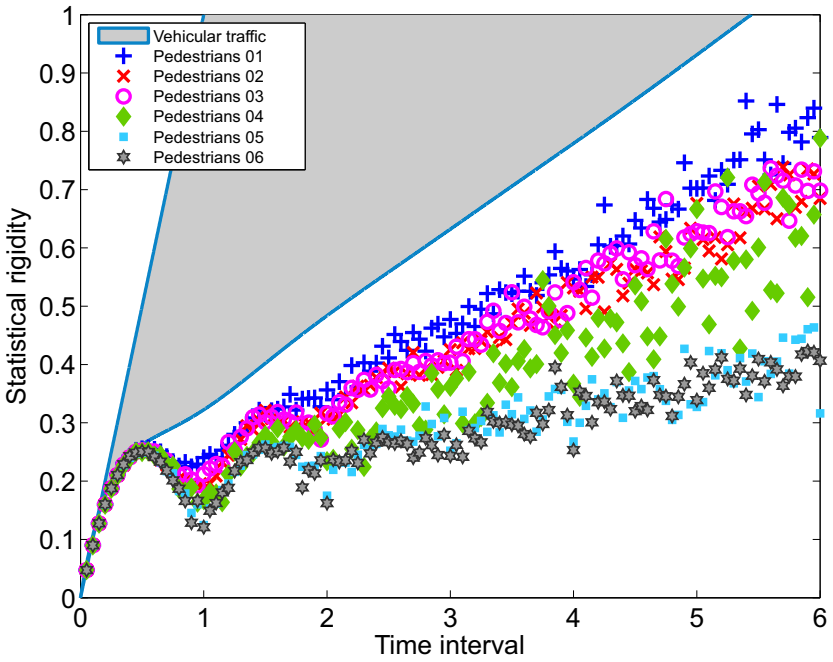

Figure 3. Graphs of the statistical rigidity analyzed for several pedestrian experiments (see [27] for details). The gray zone shows a region of vehicular rigidities, i.e. empirical statistical rigidity detected for vehicular traffic is usually a curve lying in this zone.

where μ_k. is $k-$th statistical moment associate with the generator of the BPS and $\sigma^2 = \int_{\mathbb{R}}(x-\mu_1)^2 g(x)\,\mathrm{d}x$ is the variance.

6. Classification of random systems according to compressibility

Basically, for typical theoretical or empirical systems (energy levels in quantum chaotic systems, eigenvalues of random matrices, spring chain, movement of a crowd in the corridor, Dyson's Coulomb gas, movement of vehicles on a single lane) three categories of stochastic behavior were detected. The first category consists of the Poisson systems (PS) corresponding to ensembles of uncorrelated compressible levels. In such ensembles (after being scaled to $\mu_1 = 1$), the statistical rigidity has a purely linear course $\Delta(L) = L$ and its graph creates a natural upper bound for all one-dimensional stochastic systems investigated so far, because Poisson system is an ensemble of absolutely uncorrelated events, whose fluctuations are therefore the most intense. Compressibility of the system is equal to one. The lower boundary is defined by the statistical rigidity of deterministic systems (DS), which are absolutely rigid. An example of such a system is the one-dimensional system of electrically-charged balls on a circle, whose equidistant particle arrangement does not allow any statistical deviations. It leads to zero statistical rigidity and compressibility $\chi = 0$. In RMT interpretation, the zero-rigidity states correspond to a system of rigid, incompressible levels (like for quantum harmonic oscillator).

Statistics in many other systems experience a crossover from DS to PS and are located between the two boundary curves (see figure 4), i.e. value of compressibility lies between 0 and 1. We refer to these systems as *sub-Poisson systems*. These systems include the well-known ensembles GUE/GOE of unitary/orthogonal random matrices, but also systems of pedestrians moving in a narrow corridor [27] or a set of vehicles moving on a single lane [6].

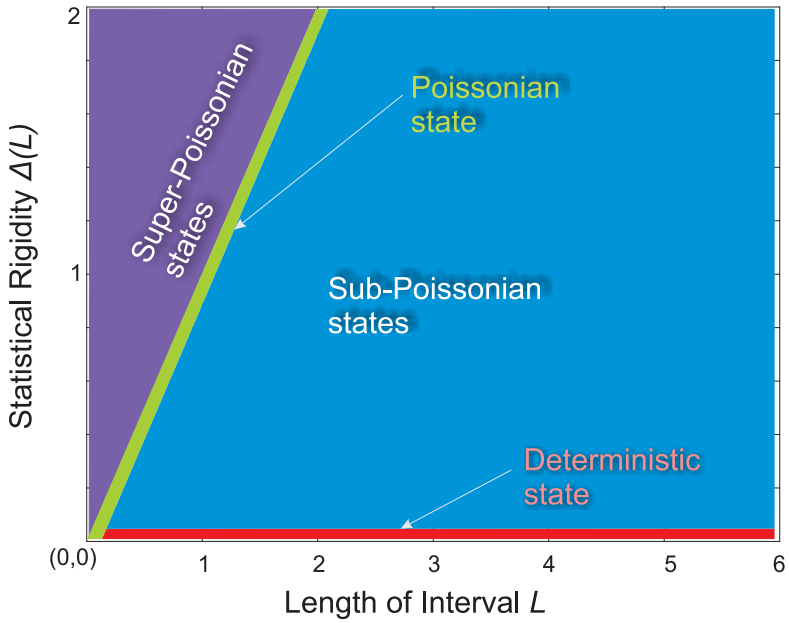

Figure 4. Classification of states in stochastic systems according to the statistical rigidity.

In a simplistic formulation, i.e. for iBPS with scaled headways, this division can be differentiated as follows: Whereas in the Poisson systems the standard deviation of particles headways is equal to the mean value, in the deterministic systems it is zero. In sub-Poisson systems, this value lies between zero and one.

From the implication (13), which can be applied under the condition (12), it directly follows that variance of scaled headways is less than one. It means that all particle systems with a purely repulsive interaction (i.e. $F(x) = -\varphi'(x) > 0$) where, in addition, $\beta > 0$ are sub-Poissonian, since formula (14) implies that for the associate compressibility it holds $\chi = \mathtt{VAR}(\mathcal{X}) = \sigma^2 = \mu_2 - 1 < 1$. Thus, for creating a super-Poisson ensemble with $\chi > 1$, particle interaction must also include an attractive component.

7. Detection of super-Poissonian states in vehicular flow

7.1. Super-Poisson systems

A system or state of a system with compressibility greater than one is considered to be an *super-Poissonian.* Compared to other types of systems, the occurrence of the super-Poissonian systems is very infrequent. Actually, many physical systems can never be in the super-Poissonian phase. A representative example of such system is the above-mentioned thermodynamic system of point-like particles with a short-ranged repulsion between succeeding particles. Here, the inter-particle headways $\mathcal{R}_0, \mathcal{R}_1, \mathcal{R}_2, \ldots$ are i.i.d. random variables, which corresponds to the theoretic concept of iBPC. Also the conditions from the section 5.2 are fulfilled. It means that the compressibility χ is equal to the variance $\mathtt{VAR}(\mathcal{R})$ and $\mathtt{VAR}(\mathcal{R}) \leqslant 1$ according to (13).

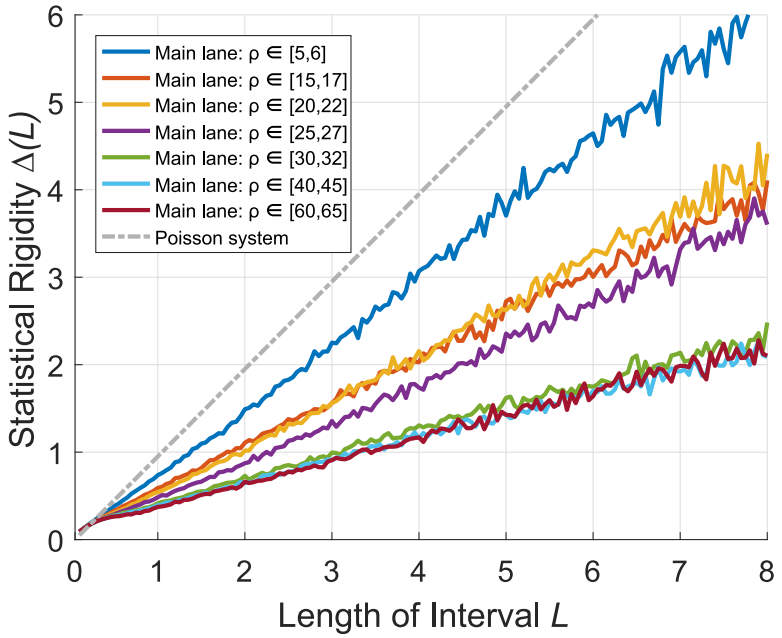

Figure 5. Statistical rigidity of freeway data for the selected segments $\Omega_{\mathtt{VD}}$ (see below) in a main lane. Vehicle-by-vehicle data have been collected during 90 days at freeway circuit R1 (located near Prague, Czech Republic) by technology of induction double-loop detection. Specifically: a) $\Omega_{\mathtt{VD}} = [5,6] \times [100,115]$, b) $\Omega_{\mathtt{VD}} = [15,17] \times [100,115]$, c) $\Omega_{\mathtt{VD}} = [20,22] \times [90,105]$, d) $\Omega_{\mathtt{VD}} = [25,27] \times [80,100]$, e) $\Omega_{\mathtt{VD}} = [30,32] \times [10,80]$, f) $\Omega_{\mathtt{VD}} = [40,45] \times [10,80]$, g) $\Omega_{\mathtt{VD}} = [60,65] \times [10,80]$. [units: veh/km × km/h]

Super-Poisson systems, whose statistical fluctuations exceed fluctuations observed in systems with totally independent/uncorrelated events, are therefore very rarely discussed in the literature.

7.2. Statistical rigidity of vehicular data samples

Statistical analysis of empirical traffic data, i.e. a course of statistical rigidity detected for specific segments, confirms that states of traffic flow in a main lane remains in the sub-Poisson territory – see figure 5. However, much more interesting behavior is detected when analyzing fast-lane data. Whereas some fast-lane samples lie in the territory of sub-Poisson states, as well, the statistical rigidity of other samples surprisingly intersects the territory of super-Poisson states. The latter represents data samples measured for lower traffic densities, when the traffic is in the free or transition phases – see figure 6.

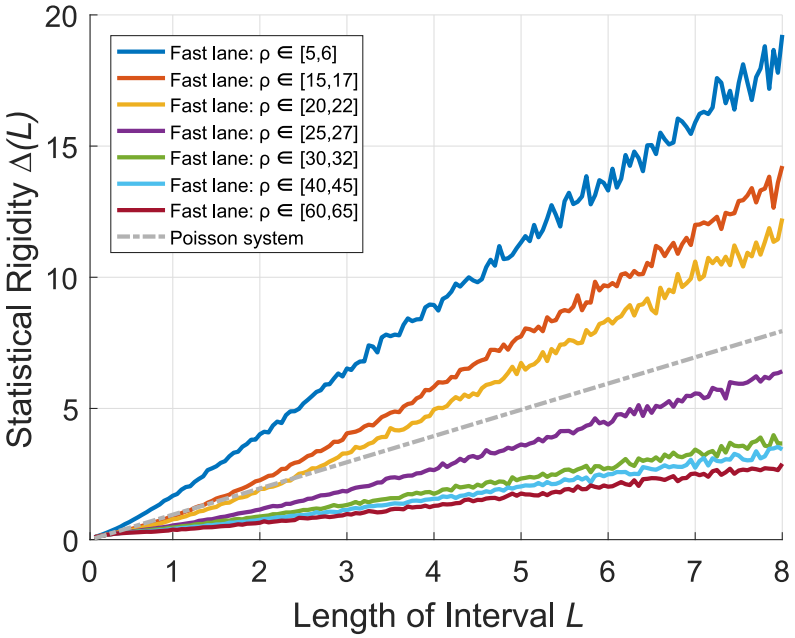

Figure 6. Statistical rigidity of freeway data in a fast-lane traffic flow. Data visualized have been extracted from the same segments as those mentioned in capture of figure 5.

More detailed analysis of the real-road rigidity has been carried out (according to the 3s-segmentation procedure) as follows. First of all, a fixed segment Ψ has been defined as a density band $\Psi = [\varrho, \varrho{+}5\,\mathrm{veh/km}]$, where the traffic density ϱ is a parameter. From the entire record it has been then extracted data sub-samples (with the sampling size $\mathtt{K} = 50$) belonging to the given segment Ψ. For the resulting $\Psi-$adjoint set of traffic micro-quantities (clearances, headways, velocities, and so on) it has been calculated a course of statistical rigidity, which is illustratively shown in figures 5 and 6. With help of robust regression methods, eliminating the curvilinear behavior of rigidity near the origin, we have quantified the value of compressibility χ. Changing the traffic density ϱ in a segment Ψ we have obtained a dependency $\chi = \chi(\varrho)$ describing of how rigid the traffic streams is. Results of this analysis are plotted in figure 7. They confirm that $\chi < 1$ for all segments of a main lane. The same behavior is visible for fast-lane data extracted

from a congested traffic phase for densities larger than 25 veh/km. Contrariwise, free-flow compressibility exceeds the border value $\chi = 1$, which means that associate traffic states are super-Poissonian.

To eliminate doubts about the universality of this anomalous behavior we have analyzed several independent vehicular data-samples as well, mainly vehicle-by-vehicle data measured on the Dutch two-lane freeway A9. Indeed, respective compressibility shows (see figure 7) similar features as those described above. Furthermore, in Appendix 10.4 we show that super-random traffic conditions can also be found in traffic situations where speed of vehicles is significantly reduced by regulations prohibiting to cross a certain speed limit. Besides, analogous behavior has been observed also in [33].

8. Explanation of super-Poissonian behavior

The performed test reveals very interesting discrepancies between main-lane and fast-lane vehicular streams. However, for densities greater than 40 veh/km it can hardly be distinguished between both lanes. The reason for such behavior is following. At high densities, lane changes related to overtaking vehicles are extremely improbable. Both lanes are significantly synchronized and the flow in them is equivalent more or less. A completely different situation occurs for flow at lower densities, where significant freedom for drivers' decisions causes inhomogeneous division of vehicles into two lanes. While drivers, who prefer slower driving, do not change lane, other drivers take the opportunity to overtake slower vehicles and shift to a fast lane. This situation leads to more rigid arrangement of vehicles in a main lane reflecting stronger orderliness of vehicles. On the other side, in a fast lane one can recognize behavior whose stochastic level is more intense than for absolutely stochastic system of randomly distributed particles – Poissonian system. It means that these traffic states are super-random.

8.1. Explanation based on hypothesis of independence of traffic flows in different traffic lanes

Under the simplistic assumption that the flows in both lanes are independent, the above-discussed difference between the statistical rigidities measured in the two lanes can be explained as follows. In view of the theoretically proven fact that in particle systems with a purely repulsive force component (see the last paragraph of section 6), the compressibility is undoubtedly less than or equal to one, it can be argued with certainty that the interaction between fast-lane vehicles contains an attractive component in addition to the commonly-used repulsive component. This follows directly from the considerations discussed in the previous text.

Above that, in Appendix 10.2 it is shown that traffic microstructure of two dependent flows (in the main and fast lanes) is the same as in particle systems, whose mutual interactions are composed of attractions and repulsions.

Therefore, being inspired by outputs of the rigidity test we suggest the dynamic

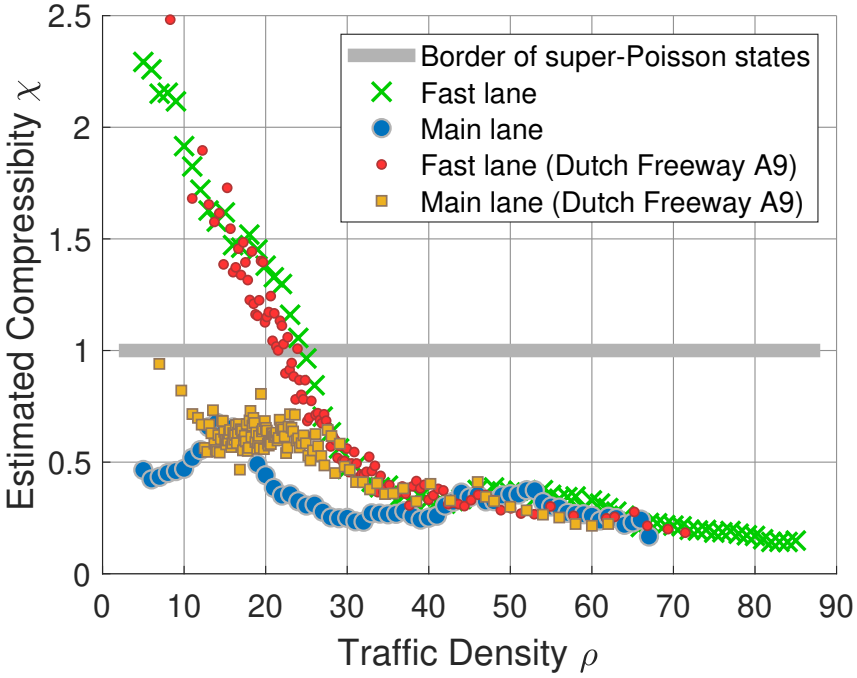

Figure 7. Compressibility of Czech and Dutch freeway data.

description combining the both, attractive and repulsive components. To be specific, a 2-component force/potential proposed may be written as

$$F(x) = -\frac{\varkappa}{x} + \frac{1}{x^2}, \quad \varphi(x) = \varkappa \ln(x) + \frac{1}{x}, \quad x > 0, \tag{15}$$

respectively, where $\varkappa \geqslant 0$ is the constant expressing a ration between repulsive and attractive components. Thus, the potentials used commonly in the transportation literature [7, 8, 6, 30, 31, 5] represent, in fact, special purely repulsive variants of (15), where $\varkappa = 0$.

The above-suggested force description implies the clearance distribution, which is described mathematically via two-parametric family: Generalized Inverse Gaussian distribution – GIG [32]

$$g_{\alpha,\beta}(x) = Ax^{-\alpha}\mathrm{e}^{-\frac{\beta}{x}}\mathrm{e}^{-\lambda x} \quad (x > 0), \tag{16}$$

where $\alpha \geqslant 0$, $\beta \geqslant 0$ represent distribution parameters. Positive constants $A = A(\alpha, \beta)$, $\lambda = \lambda(\alpha, \beta)$ ensure the proper normalization and scaling, i.e. $\mu_0(g) = \mu_1(g) = 1$. Parametr α is connected to a force constant $\varkappa$ via $\alpha = \varkappa \cdot \beta$. Therefore, for purely repulsive potentials it holds $\alpha = 0$, whereas for potentials with a significant attractive component it holds $\alpha > 0$.

A procedure estimating values α, β from empirical traffic samples is based on standard Minimum Distance Estimation method. It means that we minimize the L_2-distance

$$\sigma(g_{\alpha,\beta}, h) = \|g_{\alpha,\beta} - h\| = \left(\int_0^\infty (g_{\alpha,\beta}(x) - h(x))^2 \,\mathrm{d}x\right)^{1/2}$$

between the empirical histogram-function $h(x)$ and p.d.f. (16), i.e. we solve a optimization problem

$$(\hat{\alpha}, \hat{\beta}) = \text{argmin}_{\,\alpha\in\mathbb{R},\beta\geqslant 0}\ \sigma(g_{\alpha,\beta}, h).$$

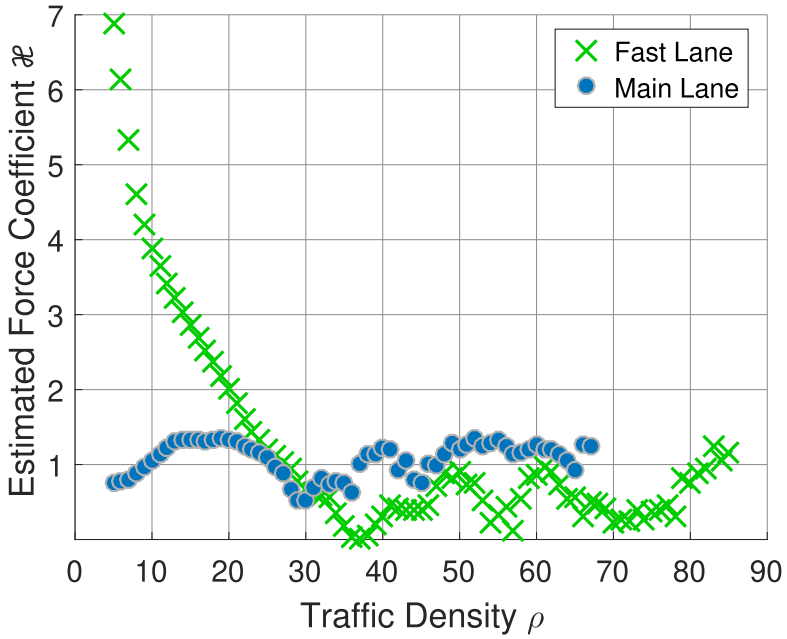

Figure 8. Estimated value of the force coefficient $\varkappa$ extracted from empirical traffic samples by means of Minimum Distance Estimation method. Estimation procedure has been applied to distance-clearance distributions analyzed in the same segments as discussed in section 7.

Estimated values $\hat{\beta}$ and $\hat{\varkappa}$ are visualized in figures. 8 and 9. As apparent, statistical resistivity of segmented samples shows almost linear increase with traffic density, which means that statistical self-organization of vehicular systems grows from slightly correlated states being close to non-correlated events in Poissonian systems to strongly organized states, which converge (for resistivity approaching to infinity) to absolutely rigid deterministic systems. The behavior of the second parameter show considerably more interesting features. For congested traffic regime (with traffic density above 30 veh/km) the force coefficient $\varkappa$ gains very low values, which means that a leading interaction term is a repulsion between succeeding cars, which holds similarly for both lanes. However, for free-flow and transition regimes there is a significant difference between fast and slow lanes. Whereas main-lane vehicles are attracted to their forerunners quite weakly, for fast-lane drivers the influence of a attractive force component prevails over a repulsive component. Such a difference between both lanes clearly distinguishes a competitive way of driving in fast lanes from tranquil maneuvering of cars in a main lane. It is also very clear from figure 9 that the aggressive nature of fast lane maneuvering is suppressed in the condensed phase due to the increased traffic

density, when an aggressive driving style may be dangerous. The increase in traffic density then visibly causes synchronization not only in driving style, where the ratio between attractive and repulsive force component is practically the same in both lanes, but also in the stochastic synchronization in vehicular microstructure as the respective headway/clearance distributions are very similar in both lanes.

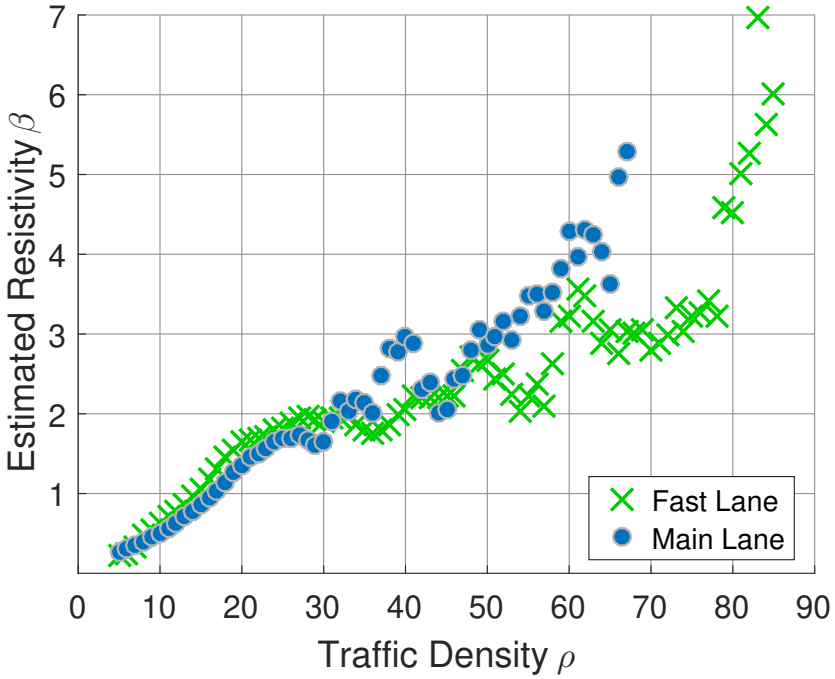

Figure 9. Estimated value of statistical resistivity β extracted from empirical traffic samples by means of Minimum Distance Estimation method.

Such an argumentation has, however, two substantial drawbacks. The premise of independence of traffic streams in the fast and slow lane is at least debatable. In fact, the both streams are strongly dependent, predominantly for intermediate traffic densities. Moreover, the theoretical assumption of statistical independence of succeeding headways/clearances, which was used in the theoretical concept of the balanced particle system, is not reasonable in some cases [21]. Therefore, it is necessary to consider the foregoing considerations as illustrative only. However, they will help us to comprehend why certain traffic samples show a higher rate of fluctuations than systems in which the values of selected quantities are completely random.

8.2. Explanation based on concept of vehicular bunching

Against the expectation, the statistical compressibility gained for the low-density fast-lane traffic samples – see figure 7 – takes on values higher than one. It makes the associate state super-Poissonian. Analogously, super-Poissonian distributions have been found for example in photon counting experiments and statistics of light, in quantum optics [13, 14, 15]. In these photon experiments, it has been revealed a statistical

tendency of photons to arrive simultaneously at a detector, which results in fact, that the corresponding probability distribution of photon arrival time at the detector is super-Poissonian. This phenomenon, called *photon bunching,* is attributed to the fact that photons are Bosonic. In physics it is known also as the Hanbury Brown and Twiss effect [16].

Super-Poissonian behavior of vehicular traffic, revealed only in multi-lane traffic data-samples, originates in the drivers lane-changing behaviour causing *a vehicular bunching* in a fast lane of a freeway. Imagine a common traffic situation experienced on freeways. A driver in the slow lane intends to overtake a slower vehicle and starts to change to the fast lane. When there is a vehicle moving in the fast lane approaching to the lane-changing vehicle, it might need to slow down in order to avoid collision. When there are more vehicles moving in the fast lane, they are slowed as well, which forms a bunch of vehicles moving within a close distance. The bunching has the same physical effect as the presence of attractive forces in the particle systems discussed above. Moreover, not many of such overtaking manoeuvres originating in the slow lane are necessary to block the fast lane locally, which results in traffic intensity fluctuations in the fast lane. As a consequence, the fast lane exhibits super-Poissonian like statistics for traffic densities up to 25 vehicles per kilometer. To support these assertions we show in figure 10, how a rate of fast-lane vehicles, whose speed exceeds the permitted limit 130 km/h, depends on traffic density. Indeed, for densities above 25 vehicles per kilometer this rate is essentially zero.

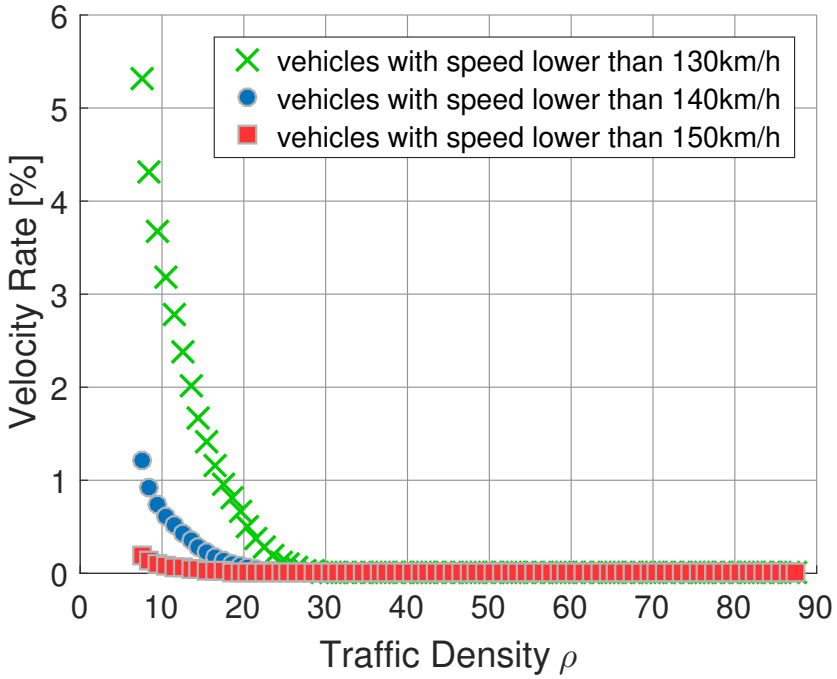

Figure 10. Density-dependent rate of vehicles exceeding the permitted speed limit (and other limits) for a fast freeway-lane.

An interesting parallel with super-Poissonian states on freeway can be found in the field of traffic safety, as well. Here the similar effect is known as over-dispersion and it used to be modelled by means of the mixture of Poisson distributions whose single parameter is distributed according to two-parametric Gamma distribution. This leads to the negative binomial distribution describing crash-frequencies [34, 35].

9. Summary

For this paper the leveling and well-researched system is the so-called Poisson system, i.e. a one-dimensional particle system whose occurrences of particles on an interval of a given length are subject to a well-known Poisson distribution. This system is, in fact, a system with an absolute degree of randomness, because the locations of its particles are chosen completely at random, without any correlation between positions of individual particles. Therefore, this system is usually understood as a system with the maximum rate of statistical fluctuations.

For the purposes of this paper, these fluctuations are quantitatively described using a well-established tool called statistical rigidity, which represents the statistical variance of the number of particles occurring on the interval of length L. The statistical rigidity in Poissonian systems has a purely linear course $\Delta(L) = L$. Most known physical, biological, socio-dynamic systems, however, show a lower rate of fluctuations, when the slope χ of the linear asymptote for statistical rigidity $\Delta(L) \approx \chi L + \mu$ is less than one (sub-Poisson systems). The slope χ, representing the derivative of statistical rigidity performed in the region of large L, is referred to as statistical compressibility. With the help of the instrument of compressibility, it is then possible to straightforwardly classify individual states of a one-dimensional stochastic system. For $\chi = 1$ it is a Poisson system, for $0 < \chi < 1$ it is a sub-Poisson system. A zero compressibility indicates a purely rigid system without any fluctuations.

However, in this article the main objects of the analyzes are the so-called super-Poisson states. These are the states, in which the statistical fluctuations of selected statistical quantities are higher than fluctuations in the Poisson system. The compressibility of such a system is therefore greater than one, i.e. the rigidity graph rises steeper than 45 degrees. Super-Poisson states have been revealed by analyzing vehicle-by-vehicle data measured in a fast lane at densities not exceeding 25 vehicles per kilometer. In other cases, the associate traffic states are sub-Poissonian.

Empirical explanation for the formation of super-random states is based on an analogy with photon counting experiments, where the presence of super-random states has been detected, as well. As in the theory of Hanbury Brown and Twiss effect, we also explain the origin of these states using a bunching. Vehicular bunching (according to our interpretation) occurs when vehicles in the fast lane are forced to brake due to a vehicle currently moving from the slow to the fast lane. This situation leads to the formation of local clusters, which – from a global perspective – represent an anomalous subgroup of vehicles. The presence of these local clusters then leads to an increase in

fluctuations on a global scale. However, these effects cease as soon as the increased traffic density prevents overtaking.

In the article, this argumentation is justified by a simple traffic model, in which a steady-state arrangement of vehicles in the fast lane is disrupted by overtaking maneuvers of vehicles from the slow lane. Indeed, the results of our simulations confirm that the inclusion of main-lane vehicles in the fast-lane flow causes a transition of the system from the sub-Poisson phase to the super-Poisson phase. Furthermore, it is shown that these inclusions have the same effect as if they were present also attractive stimuli between elements in the system. It means that traffic microstructure resulting from overtaking is identical with the microstructure, which is spontaneously formed in a stochastic particle-system with a force-description containing, in addition to a traditional repulsive force-component preventing vehicle collisions, also an attractive component. Its existence may be due either to the competitiveness of drivers, when more sporty drivers try to compete with others, or/and it may arise indirectly during overtaking maneuvers, where vehicles in a faster lane are approaching the vehicles that have just completed an overtaking maneuver and do not yet have the appropriate speed. Statistical detection, based on standard estimation procedures applied to clearance distributions, made it possible to estimate the degree of these attractive forces during changing traffic conditions. In accordance with an intuitive idea, it has indeed been shown (see figure 8) that a more significant attractive component is present only in the fast lane as far as current traffic parameters allow vehicles to overtake. As soon as the traffic density reaches a critical value (here approximately 25 veh/km), driving in both lanes is significantly synchronized and overtaking processes cease. At the same time, the considered attractive force-component also disappears.

All the above argumentations are supported by the presented results of the statistical analyzes. In addition to the results already mentioned, it is also worth noting that the detected statistical resistivity β (i.e. stochastic-noise resistance) has a very similar course in both lanes (see figure 9), which means that (in contrast with attractive force component in figure 8), the noise level in both lanes is practically the same. In addition, it is also necessary to mention that super-random states do not occur on one-way single-lane roads, which fully supports the above-mentioned explanation.

Acknowledgments

Research presented in this work has been supported by the Grant SGS18/188/OHK4/-3T/14 provided by the Ministry of Education, Youth, and Sports of the Czech Republic (MŠMT ČR). The authors would also like to thank The Road and Motorway Directorate of the Czech Republic (Ředitelství silnic a dálnic ČR) for providing traffic data analyzed in this paper.

References

[1] Treiber, M., Kesting, A., 2013. Traffic Flow Dynamics, Berlin: Springer.

[2] Kerner, B.S., 2004. The Physics of Traffic, Springer-Verlag, New York.
[3] Helbing, D., Treiber, M., 2003. Phys. Rev. E 68, 067101.
[4] Treiber, M., Kesting, A., Helbing, D., 2006. Physical Review E 74, 016123.
[5] Mahnke, R., Kaupuzs, J., Hinkel, J., Weber, H., 2007. The European Physical Journal B 57 (4), 463.
[6] Krbálek, M., Šeba, P., 2009. J. Phys. A: Math. Theor. 42, 345001.
[7] Krbálek, M., 2007. J. Phys. A: Math. Theor. 40, 5813.
[8] Krbálek, M., Helbing, D., 2004. Physica A 333, 370.
[9] Krbálek, M., Hobza, T., 2016. Physics Letters A 380 (21), 1839.
[10] Kollert, O., Krbálek, M., Hobza, T., Krbálková, M., 2019. Journal of Physics Communications 3, 035020.
[11] Krbálek, M., Šleis, J., 2015. J. Phys. A: Math. Theor. 48, 015101.
[12] Mehta, M.L., 2004. Random matrices (Third Edition). New York: Academic Press.
[13] Morgan, B. L., Mandel, L., 1966. Phys. Rev. Lett., V. 16, 1012.
[14] Nakayama, K., Yoshikawa Y., Matsumoto, H., Torii, Y., Kuga, T., 2010. Optics Express, V. 18, 6604.
[15] Park, J., Jeong, T., Moon., J. S., 2018. Scientific Reports, V. 8, 10981.
[16] Hanbury Brown, T., Twiss, R. Q., 1957. Proceedings of the Royal Society, Series A, Mathematical and Physical Sciences, 242 (1230), 300.
[17] Cowan, R.J., 1975. Transp. Res. 9, 371.
[18] Krbálek, M., 2013. J. Phys. A: Math. Theor. 46, 445101.
[19] Treiber, M., Helbing, D., 2009. Eur. Phys. J. B 68, 607.
[20] Kumm, M., Schreckenberg, M., 2020. Collective Dynamics, 5, A81: 1-17
[21] Krbálek, M., Apeltauer, J., Apeltauer, T., Szabová, Z., 2018. Physica A 491, 112.
[22] Helbing, D., 2001. Rev. Mod. Phys. 73, 1067.
[23] Kerner, B.S., Rehborn, H., 1996. Phys. Rev. E 53, R4275.
[24] Neubert, L., Santen, L., Schadschneider, A., Schreckenberg, M., 1999. Phys. Rev. E 60, 6480.
[25] Chalker, J.T., Kravtsov, V.E., Lerner, I.V., 1996. Jetp Lett. 64, 386.
[26] Bohigas, O., 1991. Random matrix theories and chaotic dynamics (IPNO-TH–90-84). France
[27] Krbálek, M., Hrabák, P., Bukáček, M., 2018. Physica A 490, 38.
[28] Cox, D. R., 1962. Renewal Theory, Methuen & Co., London, New York.
[29] Cox, D. R., Lewis, P. A. W., 1978. The statistical Analysis of Series of Events, Chapman & Hall.
[30] Treiber, M., Hennecke, A., Helbing, D., 2000. Physical Review E, 62 (2), 1805.
[31] Gazis, D.C., Herman, R., Rothery, R. W., 1961. Operations Research 9, 545.
[32] Jörgensen, B., 1982. Statistical Properties of the Generalized Inverse Gaussian Distribution, Lecture Notes in Statistics, 9, Springer, Heidelberg.
[33] Šeba, F., Krbálek, M., 2020. Super-Poissonian Statistics In Traffic Flow, APLIMAT 2020 - Proceedings, 930.
[34] Miaou, S.P., Lord, D., 2003. Transport. Res. Rec. 1840, 31.
[35] Mitra, S., Washington, S., 2007. Accident Analysis and Prevention, 39, 459.

10. Appendix

10.1. Stochastic many-particle gas – physical realization of the balance particle system

The physical realization of the above-discussed balanced particle system (see subsection 5.3) is represented by an ensemble of many identical particles located along a curve (typically a semi-line, line, or circle) subjected to stochastic noise of various intensity. The noise level in this system is controlled by the so-called *stochastic resistivity* coefficient β. The zero value of β implies an absolute noise level in the ensemble,

which therefore corresponds to the Poissonian system where purely random locations of particles generate the exponentially distributed inter-particle headways, whose steady-state headway distribution reads

$$g(x) = \Theta(x)\lambda \mathsf{e}^{-\lambda x}.$$

Here $\mathtt{E}(X) = 1/\lambda$. Conversely, if the resistivity β is increasing above all limits the system corresponds in fact to a deterministic system, whose inner dynamics is not burdened with any stochastic fluctuations. Under this condition it holds $g(x) = \delta(x - \nu)$, which means that particles are located equidistantly in locations $x_0, x_0+\nu, x_0+2\nu, x_0+3\nu, \ldots$. Here $\delta(x)$ stands for the Dirac delta.

Arrangement of particles in the two border variants does not depend in any way on interactions between individual particles. However, the situation will dramatically change for intermediate values $0 < \beta < +\infty$. Then the stationary state of the system will be strongly dependent not only on the value of resistivity, but especially on interaction forces that determine mutual interactions between neighboring particles in the system. In a homogeneous variant of the system, when all neighboring particles are repulsed/atracted via the same force description (force $F(x)$ and interaction potential $\varphi(x)$), where $F(x) = -\varphi'(x)$, the associate steady-state of the system is described (see general derivation in [7]) by the following inter-particle headway distribution:

$$g(x) = A\Theta(x)\mathsf{e}^{-\beta\varphi(x)}\mathsf{e}^{-\lambda x}, \tag{17}$$

where constants $A = A(\beta)$ and $\lambda = \lambda(\beta)$ ensure the proper normalization and scaling.

In addition, from this general derivation and from the considerations presented at the end of the section 6 one can assert that for creating a super-Poisson ensemble (having compressibility greater than one) particle interaction must also include an attractive component.

10.2. Vehicular bunching and super-Poissonian traffic states – two consequences of overtaking

In this part of Appendix we aim to illustrate that an overtaking of vehicles may cause vehicular bunching, which is responsible for creation of super-Poissonian states in traffic microstructure. Moreover, we try to show that inclusion of main-lane vehicles in a gap between fast-lane vehicles acts effectively as a kind of a attractive force present in fast-lane traffic stream. Our considerations are based on the idea that overtaking maneuver brings a significant local instability into the steady-state arrangement of vehicles in the fast lane, which finally causes the resulting atypical arrangement of vehicular positions, previously referred to as super-Poisson. To support our assumption, we have constructed the following simulator based on the above-discussed stochastic gas and simulated by the modified Metropolis algorithm [11] having the following attributes: force-potential $\varphi(x)$, stochastic resistivity β, critical gap γ, and acceptance ratio q (see explanation below).

For the afore-mentioned purposes we use the following well-known findings. Firstly, overtaking maneuvers are frequent only at lower traffic densities, when one can find more fast-lane gaps suitable for inclusion of a main-lane vehicle. Therefore, we will consider a low value of stochastic resistivity β in our model, which corresponds (as can be clearly seen in figure 9) to traffic systems with lower density, that admits higher statistical fluctuations. Secondly, as proven by many researches [18, 11, 7, 8, 6, 30] one of the most suitable potentials utilized for vehicular headway modeling is the hyperbolic potential $\varphi(x) = 1/x$. Indeed, this choice leads (after applying formula (17)) to the GIG-distributed headways, which is in an excellent compliance with empirical microstructure of traffic [7, 8, 11, 18]. Thirdly, an overtaking maneuver can be realized only if there is a gap between fast-lane vehicles greater than a certain critical value (critical gap γ). However, a sufficiently large gap can sometimes remain unused, but sometimes it can be utilized to absorb an overtaking vehicle. If the latter occurs, the long gap is divided into two randomly-long parts. It means that the maneuver will take place only with a certain probability q (acceptance ratio).

The basic concept of the model is as follows. Particles randomly-moving on a abscissa (respecting an original order and fulfilling periodic boundary-conditions) are mutually repulsed through the potential $\varphi(x)$ and randomized by a noise with intensity driven by β. After reaching a steady state, inter-particle headways greater than critical gap γ are ready for possible overtaking procedure. With probability q such a headway is randomly divided into two segments. Otherwise, i.e. with probability $1 - q$, the headway remains unchanged. Thereafter, the resulting set of headways obtained from the simulation (and scaled to the unit average) is then subjected to standard statistical analysis.

Outputs of the analysis show that for appropriately-chosen parameters, when $q \gg 0$, (see caption below figure 11) one can acquire headway-sets having features typical for super-Poissonian ensembles. Indeed, as visible in figure 11, variance of numerical data visualized (crosses) is significantly greater than one, which indicates a super-Poisson state of bunched particles. In addition, probability density GIG (16) – obtained by statistical estimations based on MLE approaches – fits the numerical data quite convincingly. It means that the same distribution of inter-particle headways (as that detected in simulated data) can be gained from a steady state of the particle gas with a potential combining repulsive and attractive stimuli. To be specific, optimal GIG-estimation (16) of numerical data visualized by crosses in figure 11, where $\hat{\alpha} = 1.1489$, $\hat{\beta} = 0.2018$, $\hat{\lambda} = 0.2906$, represents a steady-state distribution of a gas with resistivity $\hat{\beta}$ and potential

$$\varphi(x) = \frac{\hat{\alpha}}{\hat{\beta}} \ln(x) + \frac{1}{x}.$$

It means that traffic microstructure resulting from overtaking is identical with the microstructure, which is formed in a gas with a potential containing, in addition to a repulsive stimulus, also an attractive component.

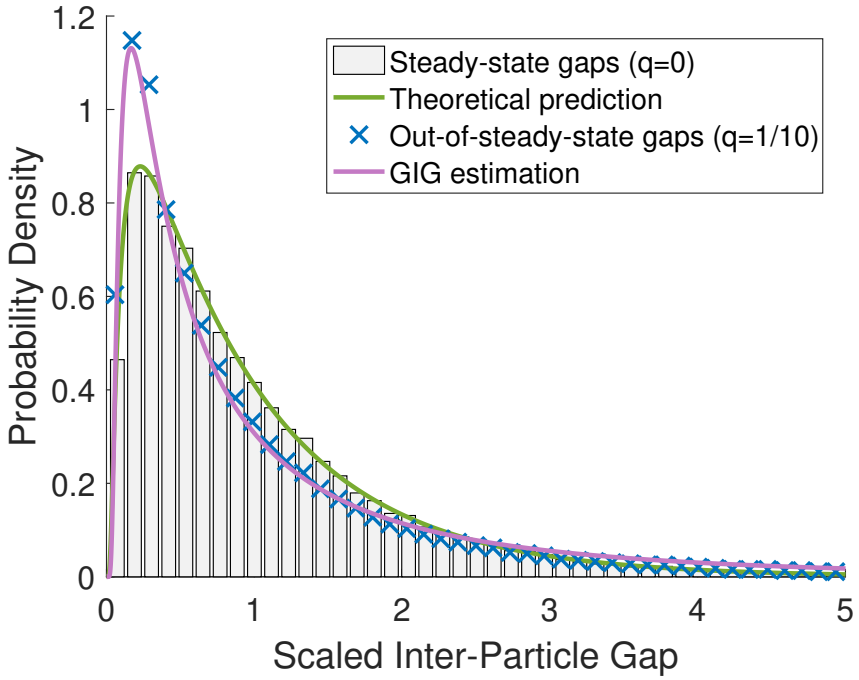

Figure 11. Distributions of scaled gaps in the numerical model. We visualize two simulation outputs for the following parameters. Gray bars: $\beta = 1/10$, $\gamma = 1/3$, $q = 0$. Blue crosses: $\beta = 1/10$, $\gamma = 1/3$, $q = 1/10$. Green curve shows the analytically-derived prediction (17), which is valid for system without overtaking. Purple curve has been acquired by Maximum Likelihood Estimation assuming that synthetical gaps are GIG-distributed. Compressibility of steady-state data is $\chi = 0.8261$ (sub-Poissonian state), whereas compressibility of perturbed data is $\chi = 1.0561$ (super-Poissonian state).

10.3. Cross-validation and statistical reliability

In order to demonstrate the consistency and statistical reliability of the results presented we test here whether super-Poisson states are present also in other vehicle-by-vehicle data and also in smaller subsets of the main data file used in this paper. To be specific, we apply the procedure estimating a sample compressibility χ repeatedly for four disjoint subsets taken from different segments of highway circuit R1 Prague, Czech Republic. Above that we make similar estimations for four disjoint data-files measured on the Dutch two-lane freeway A9. The results of the estimations are plotted in figures 12, 13. In all cases the compressibility of lower-density traffic states measured in a fast lane exceeds (in contrast to main-lane data) the limit defined by the pure Poisson states. Essentially the same behavior is detected also in [33], where the similar procedures have been applied to public traffic data.

10.4. Super-random states in urban traffic

In the last part of the text, we would like to show that super-random traffic conditions can also be found in traffic situations where speed of vehicles is significantly reduced

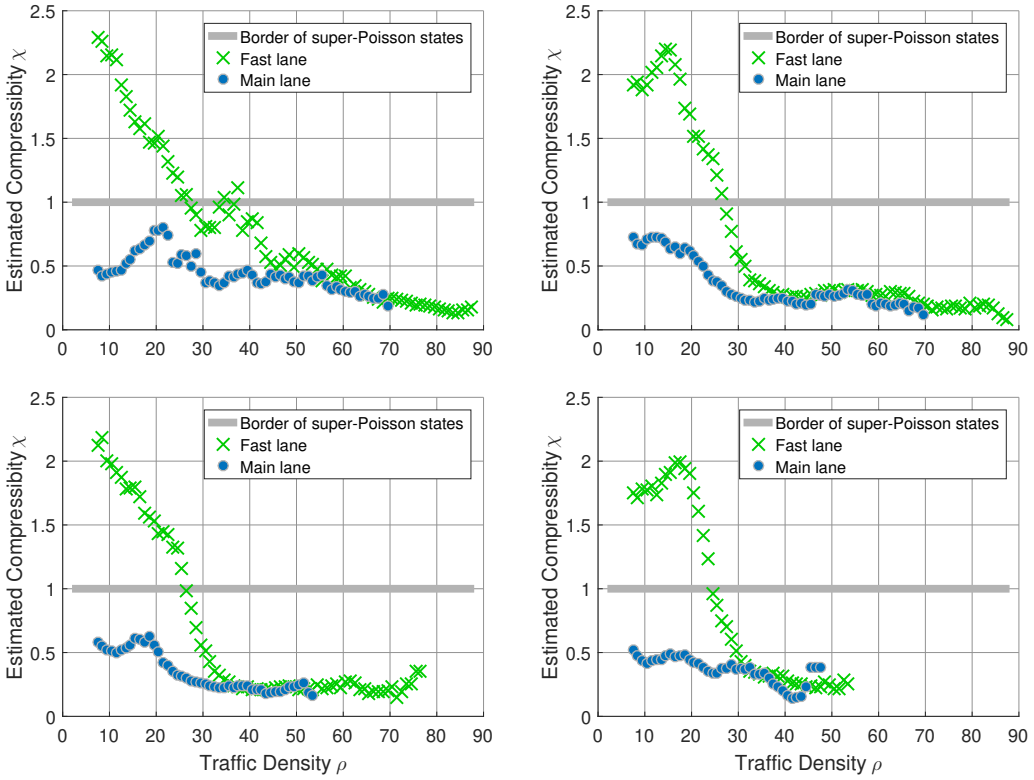

Figure 12. Compressibility of empirical vehicular data (Czech Republic). We plot estimated values χ for 4 separate independent subsets of the data-file analyzed in this paper. In regions with insufficient amount of data, i.e. typically for very-low density regions or for density above ≈ 60 veh/km, the compressibility is not ascertained.

by traffic signs. Such a situation typically occurs within urban agglomerations, where relevant regulations prohibit crossing a certain speed limit (in the Czech Republic it is 50 kilometers per hour). However, more sporty drivers usually utilize the fast lane to overtake main-lane vehicles that respect the speed limit. In this way, the super-random states are generated in urban traffic patterns as well. This is very clearly seen in figure 14 that demonstrates the results of respective statistical analysis. A curve of statistical rigidity has been extracted from about 500,000 vehicle passages detected by technology of video object detection. The linear part of the statistical rigidity has been subsequently subjected to a standard linear regression. This method has been applied to determine the empirical value of statistical compressibility, whose dependence on traffic density is plotted in the figure.

The result convincingly shows the presence of super-random states in almost all density zones of the fast lane. In contrast, all main-lane traffic states are strictly sub-Poissonian, which exactly replicates the situation found for expressway traffic.

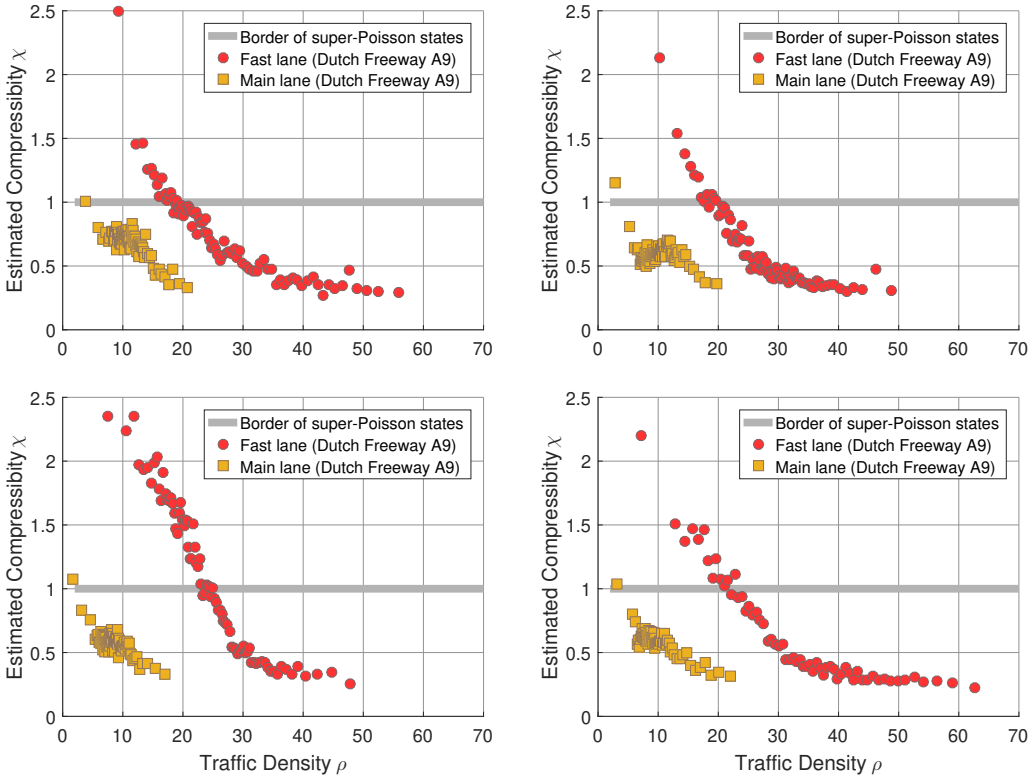

Figure 13. Compressibility of empirical vehicular data (Netherlands). We plot estimated values χ for four disjoint data-files measured on the Dutch two-lane freeway A9. In regions with insufficient amount of data, i.e. typically for very-low density regions or for density above ≈ 50 veh/km in a fast lane or ≈ 25 veh/km in a main lane, the compressibility is not ascertained.

10.5. Variable list and explanation

For comfort and better understanding, in the following table we summarize variables used in this paper and explain their meaning.

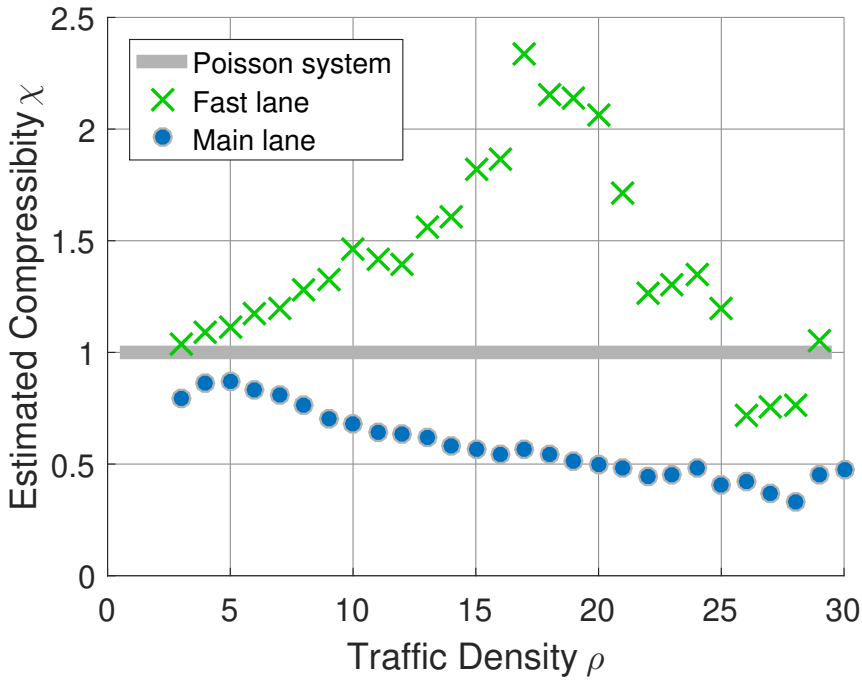

Figure 14. Estimated compressibility for urban traffic. We plot values of the statistical compressibility, which have been obtained by linear regression applied to a course of the empirical statistical rigidity. This has been extracted from urban traffic data measured in several Czech cities. Data from both lanes have been analyzed separately.

Table 1. List of variables.

Technical term	Symbol	Explanation
index	k	index of a vehicle
instant time	τ	
spatial location	ξ	
expected value and variance	$\mathtt{E}(\mathcal{X})$, $\mathtt{VAR}(\mathcal{X})$	descriptive characteristics of variable $\mathcal{X}$
ith statistical moment	μ_i	$\mu_0 = 1$, $\mu_1 = \mathtt{E}(\mathcal{X})$, and $\mu_2 = \mathtt{VAR}(\mathcal{X}) + \mu_1^2$
stochastic resistivity	β	parameter in stochastic particle gas
sampling size	$\mathtt{K}$	we use $\mathtt{K} = 50$ in the paper
incoming time	$\tau_k^{(\mathrm{in})}$	instant when a front bumper reached a detector
outgoing time	$\tau_k^{(\mathrm{out})}$	instant when a front bumper reached a detector
velocity	v_k	velocity of the kth car
vehicle length	ℓ_k	length of the kth car
incoming time container	$T^{(\mathrm{in})}$	collects time instants $\tau_k^{(\mathrm{in})}$
outgoing time container	$T^{(\mathrm{out})}$	collects time instants $\tau_k^{(\mathrm{out})}$
velocity container	Υ	collects velocities of cars
length container	Λ	collects vehicle lengths
vehicle position 1	$\xi_k^{(\mathrm{front})}$	location of a front bumper
vehicle position 2	$\xi_k^{(\mathrm{front})}$	location of a rear bumper
segment (segmentation zone)	Ψ	small sub-area v ID plane
first fundamental relation	$\Omega_{\mathtt{ID}}$	all existing intensity-density pairs
second fundamental relation	$\Omega_{\mathtt{VD}}$	all existing velocity-density pairs
time headway	z	time between two cars as they pass a detector
time clearance	t	time gap between two successive vehicles
distance headway	s	space between two front bumpers in a fixed time
distance clearance	r	space gap between two successive vehicles
scaled time clearance	y	time clearance after the unification procedure
scaled distance clearance	x	space clearance after the unification procedure
index set	G_i	all indices respective to the ith sample
sample of time headways (ith)	Z_i	set of time headways in one sample
sample of time clearances	T_i	set of time clearances in one sample
sample of space headways	S_i	set of space headways in one sample
sample of space clearances	R_i	set of space gaps in one sample
sample of velocities	Υ_i	set of velocities related to a sample
sample of scaled time clearances	Y_i	set of scaled time clearances in one sample
sample of scaled space gaps	X_i	set of scaled space gaps in one sample
sample means	$\langle T_i \rangle, \langle S_i \rangle, \ldots$	
segmented index set	F_Ψ	sample indices belonging to Ψ
segmented time clearances	Y_Ψ	set of unified time clearances belonging to Ψ
segmented space clearances	X_Ψ	set of unified space gaps belonging to Ψ
segmented velocities	Υ_Ψ	set of velocities belonging to Ψ
force	$F(s)$	force depending on headway
force potential	$\varphi(s)$	force potential depending on headway
headways and multiheadways	$\mathcal{R}_k$ and $\mathcal{X}_k$	inter-particle headways and multiheadways
statistical rigidity	$\Delta(L)$	parameterized variance for the interval frequency
statistical compressibility	χ	a slope of linear asymptote for statistical rigidity