Article

Self-Organized Criticality of Traffic Flow: Implications for Congestion Management Technologies

Jorge Laval [*]

Article

Self-Organized Criticality of Traffic Flow: Implications for Congestion Management Technologies

Jorge A. Laval

School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, United States; jorge.laval@ce.gatech.edu

Abstract: Self-organized criticality (SOC) is a celebrated paradigm from the 90's for understanding dynamical systems naturally driven to its critical point, where the power-law dynamics taking place make predictions practically impossible, such as in stock prices, earthquakes, pandemics and many other problems in science related to phase transitions. Shortly thereafter, it was realized that traffic flow might be in the SOC category, implying that conventional traffic management strategies seeking to maximize the local flows can become detrimental. This paper shows that the Kinematic Wave model with triangular fundamental diagram, and many other related traffic models, indeed exhibit SOC, thanks in part to the fractal nature of traffic exposed here on the one hand, and our need to get to our destinations as soon as possible, on the other hand. Important implications for congestion management of traffic near the critical region are discussed, such as: (i) Jam sizes obey a power-law distribution with exponent 1/2, implying that both its mean and variance become ill-defined and therefore impossible to estimate. (ii) Traffic in the critical region is chaotic in the sense that predictions becomes extremely sensitive to initial conditions. (iii) However, aggregate measures of performance such as delays and average speeds are not heavy tailed, and can be characterized exactly by different scalings of the Airy distribution, (iv) Traffic state time-space "heat maps" are self-affine fractals where the basic unit is a triangle, in the shape of the fundamental diagram, containing 3 traffic states: voids, capacity and jams. This fractal nature of traffic flow calls for analysis methods currently not used in our field.

Keywords: traffic flow; kinematic wave model; self-organized criticality; fractals; complexity; catastrophe theory; non-equilibrium critical phenomena

---

1. Introduction

There is mounting empirical evidence [1–3] suggesting the hypotheses that urban networks exhibit self-organized criticality (SOC) [4], a celebrated paradigm in the 90's that has proved very useful in countless areas of science, from earthquakes to stock prices to forest fires to astrophysics [5,6]. Despite being around for 2 or 3 decades now, the theory and control consequences of SOC–and the corresponding fractal nature of traffic flow–have not permeated the transportation engineering literature, and even in the physics literature these concepts are only acknowledged but not exploited [7–10]. This is unfortunate because *if* it turns out that urban networks do exhibit SOC, a whole new paradigm emerges for understanding urban congestion, based on the lessons learned from complexity, catastrophe theory and non-equilibrium critical phenomena [11].

But given the absence of a general theory of SOC to date [6], it is not yet possible to formally prove or disprove this hypothesis. This paper is a step towards such a theory for traffic flow. The defining characteristic of a SOC system is that it is naturally driven to be in a "critical state" where it undergoes a phase transition [12] (e.g. from liquid to solid or from free-flow to congestion). At the critical point the characteristic length scale of the system, such as the size S of forest fires or the size of traffic jams as shown here, diverges to infinity, which renders the system scale-free and thus scale invariant. This implies a power-law distribution for the jam size, whose tail probability is proportional to $s^{-\alpha}$, i.e.:

$$\Pr\{S > s\} \sim s^{-\alpha}, \tag{1}$$

with the exponent α known as the critical exponent. As opposed to the normal distribution, where it is virtually impossible to find realizations more than 3 standard deviations away from the mean value, power laws are said to have "heavy tails": there is a non-negligible probability of observing extremely large values. As a consequence, the moments of the distribution may not exist, in which case the central limit theorem and the law of large numbers no longer hold. When $\alpha > 2$, both the mean and variance are finite; when $1 < \alpha \leq 2$, as it is typically the case, the mean is finite and the variance diverges, and for $\alpha \leq 1$ both the mean and variance diverge. The case $\alpha = 1$ is popularly known as Zipf's law [13], and as it turns out $\alpha \approx 1 \pm 0.1$ for a wide variety of phenomena, from the floods of the Nile River [14] to the population of cities [15], earthquake magnitudes, allometric scaling in mammals [16], word frequencies [13], reference links on the web [17], financial returns [18], the size (number of employees, sales, assets) of companies [19,20], turbulent flows and many others [21].

SOC behavior in traffic flow models has been conjectured since the mid-nineties in the physics community [22,23]. Using ideas from the theory of percolation phase transitions and the [24] cellular automata model (NaSch model) for numerical experiments, they conjectured that (i) the critical exponent of jam durations in (1) is very close to 1/2, suggesting a relationship with the first return time of a one-dimensional random walk, and (ii) the number of vehicles under congestion scales with time as $t^{1/2}$. They conclude that SOC implies that conventional traffic management strategies seeking to maximize the flow may be detrimental as they make the system more unpredictable and hard to control. Here, we will formalize conjecture (i) using the variational theory of traffic flow [25], and point out that the scaling in (ii) is in fact $t^{2/3}$. We will also argue that current control strategies could be adapted to minimize the externalities of SOC by ensuring the network density always remains *well below* the critical density.

Surprisingly, the impact of SOC in our field has been underwhelming [26,27]. Apart from [28–30] who present some numerical evidence of SOC on networks, it appears that the subject has gone largely under the radar. One reason might be that the main insight from the early works [22,23] was that SOC in the NaSch model emerges only in "cruise-control" mode, where random perturbations are only allowed in congested traffic states. This apparent dependence on the rule details might have been interpreted as the absence of a universality class in traffic flow, implying that SOC might not be an intrinsic property of traffic flow, but just of particular models.

This paper shows that the Kinematic Wave model [31,32] with triangular fundamental diagram, and many other related traffic models including the NaSch model, indeed exhibit SOC. In our case, showing SOC comes down to showing that (i) there is critical behavior at the transition between free-flow and congestion, and (ii) the system has a tendency to be in the critical state. The paper focuses on (i) because (ii) follows immediately from our willingness to drive as fast as possible, which translates into the well-known entropy solution of the kinematic wave model where e.g. the outflow after the signal turns green is maximized precisely at the critical density.

The remainder of the paper is organized as follows. Section 2 below establishes the equivalence of foundational traffic models in the literature. Section 3 shows that a mathematical cause for criticality in traffic flow is provided by variational theory with random initial conditions around the critical density, which produces power-law statistics for jam durations and spans. Section 4 presents aggregate measures of performance such as total delay, total distance traveled, average flows and speeds, and probability of congestion. Note that these measures of performance are not heavy tailed, and are related to the area under a random walk. A discussion and outlook is provided in the last section.

2. Equivalence of Traffic Flow Models

This section recalls existing results demonstrating that a large family of well-known traffic flow models, both continuum and discrete, are fundamentally equivalent. This is important because it allows us to use existing knowledge behind each of these models to the benefit our understanding of a single theory, best known to this community as the the kinematic wave theory of [31,32] with triangular fundamental diagram.

A microscopic formulation of this theory is the celebrated Newell's car-following model [33], which can in turn be formulated as a cellular automaton (CA) [34]. The NaSch model, for which critical behavior was first conjectured, is a stochastic version of Daganzo's CA model that introduces bounded acceleration and a probability of slow down for moving vehicles.

Another important known result is that the particular values of the slopes of the fundamental diagram are irrelevant because one can always transform the output of a traffic model with a particular diagram to the output from a different diagram via simple linear transformations; see [35]. This means that, effectively, traffic flow on a single link requires no parameters (after proper normalization). Therefore, and without loss of generality, in this paper we will use an isosceles flow-density fundamental diagram, which has the virtue, among others [35], of turning Daganzo's CA model in to elementary cellular automata rule 184 [36], namely ECA 184, by using the following parameters:

$$\text{free-flow speed = wave speed = 1, jam density = 1.} \tag{2}$$

These parameters imply that both the capacity and the critical density are 1/2, that distances are measured in units of one vehicle length and that time is measured in units of 1/2 of a critical headway.

All the models in this section will be collectively referred to as "the traffic model" in what follows. The focus of this paper is on showing the critical behavior of the deterministic traffic model rather than the random NaSch model. In this way, it can be established that critical behavior does not require random traffic rules, only random initial data.

3. The Root Cause of Criticality in Traffic Flow

In this section we show that the deterministic traffic model exhibits critical dynamic behavior under random initial conditions around the critical density. This is achieved using the Variational Theory of traffic flow [25], which boils the traffic flow problem down to a minimum path problem. It turns out that around the critical density this minimum path problem is intimately related to the theory of random walks, whose return times are power-law distributed.

The basic experiment in this paper is a single-lane road segment of length L with with vehicles obeying the deterministic traffic model and starting with initial density conditions at the critical density of 1/2 plus a random noise. For the continuum traffic models the random noise is interpreted here as a *white noise* process, i.e. initial density fluctuations at any location $0 \leq x \leq L$ are assumed independent of one another and follow the standard normal distribution. For the equivalent discrete models, where the road segment is divided into L cells of jam-spacing size, the random noise becomes a Bernoulli process over the cells $i = 1, 2, \ldots L$, which can be either 0 (empty) or 1 (occupied by a vehicle) with probabilities 1-p and p, respectively, independently of other cells. Notice that p can also be interpreted as the mean density in the segment, so $p = 1/2$.

The central contribution of this paper is made possible using Variational Theory, which is based on the traffic flow surface, $N(x,t)$, that gives the cumulative vehicle number having crossed location x by time t starting from the passage of a reference vehicle. The spatial integration of the continuum and discrete processes described above are examples of the Brownian motion and random walk processes, respectively, which have played a central role in arguably all areas of science and therefore many exact analytical results are known. In our case the Brownian motion/random walk correspond to the initial conditions $N(0, x)$, which is given by the (negative) integral of the initial density. In Variational Theory, the very existence of the fundamental diagram becomes a Hamilton-Jacobi partial differential equation (PDE) [37], whose solutions can be found by solving the shortest path problem:

$$N(x,t) = \inf_{y(\cdot)} \left\{ N(y(0),0) + \int_0^t \mathcal{L}\left(y'(s)\right) \, ds \right\}. \tag{3}$$

known as the Hopf-Lax formula [38]. The fundamental diagram corresponds to the Hamiltonian of the system, whose corresponding Lagrangian is denoted $\mathcal{L}$ and gives the maximum passing rate of an

observer moving at a given speed. The shortest path problem consists in finding a trajectory $y(\cdot)$ from (x,t) to the initial (and/or boundary) data at $(y(0),0)$ that minimizes the total cost along the path, given by the term in brackets, satisfying $y(t) = x$, and $-1 \leq y'(\cdot) \leq 1$, since $y'(\cdot)$ represents a wave speed. For a homogeneous facility (where the fundamental diagram does not change) the Lagrangian becomes independent of space and time, the time integral in (3) simplifies to $t\mathcal{L}((x-y)/t)$ where y is now a scalar satisfying $x - t \leq y \leq x + t$. Accordingly, to find the solution at a generic time-space point $P \equiv (t,x)$ amounts to finding the minimum straight-line path from (x,t) to the initial boundary at $(y,0)$, where the cost of each path is $f_P(y) \equiv G(y) + t\,\mathcal{L}((x-y)/t)$, with $G(y) \equiv N(0,y)$. In the case of the isosceles fundamental diagram used here, $\mathcal{L}(v) = (1-v)/2$ and (3) simplifies to:

$$N(x,t) = \min_{x_U \leq y \leq x_D} f_P(y) \equiv G(y) + (t - x + y)/2, \tag{4}$$

where $P \equiv (x,t)$, and $x_U = x - t$, $x_D = x + t$ define the extent of the search interval; see points U and D in Figure 1. With deterministic initial conditions at the critical density we have $G(y) = (L - y)/2, 0 \leq y \leq L$, and the cost function becomes $f_P(y) = (L - x + t)/2$, which is independent of y. Therefore, *all* paths from P to the boundary are optimal (since they all have the same cost) and we let

$$N_{\text{det}}(x,t) \equiv (L - x + t)/2 \tag{5}$$

be this deterministic solution. The important point here is that this nature of the solution holds even when a white noise is added to the initial densities. In this case, $f_P(y)$ becomes a Brownian motion *with zero drift*, which revolves around a constant ($y-$independent) expected value:

$$E\{f_P(y)\} = N_{\text{det}}(x,t) \tag{6}$$

As shown next, this result implies that the lifetime of traffic jams exhibits critical behavior.
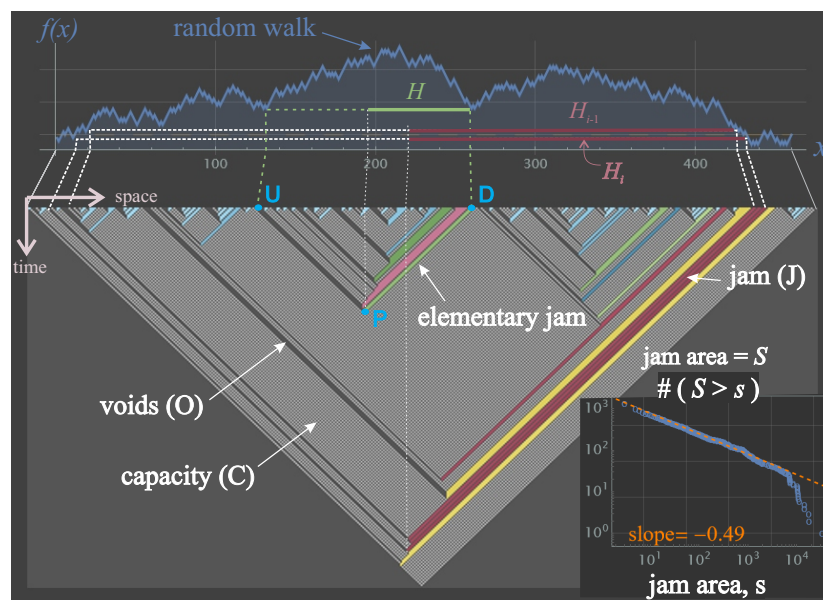
Figure 1. Relationship between the random walk/excursion $f_P(x)$ representing the fluctuations around the critical density in the initial conditions and the corresponding exact solution of the deterministic traffic flow model in the time-space diagram. The colored bars represent the congestion clusters or jams (traffic state J) in the time-space diagram, whose statistics are given by the return times of the random walk as shown. The dark gray bars are the voids (traffic state o) and the light gray areas the capacity state C. Notice that since these are the only traffic states possible, the total spatiotemporal area $A = A_0 + A_C + A_J$.

3.1. Lifetime of Elementary Jams

Let an *elementary jam* be a line segment in the (x, t)-plane that starts at a congested point D on the initial boundary and proceeds upstream at the congested wave speed so long as the visited cells are congested/occupied; see Figure 1. This is akin to a characteristic line segment connecting D to P but only passing through congested points in the plane and ending when reaching an empty cell. Ending points such as P in the figure are characterized by a transition from free-flow to congestion at that point, which means that the cost of minimum paths P-U and P-D in the figure both solve (4) and therefore have identical cost, i.e. $f_P(U) = f_P(D)$. This means that the duration $D - U$ in Figure 1 is given by the first return time of our Brownian motion $f_P(y)$, i.e. the time it takes for the Brownian motion starting at U to "return" to the same value. These return times are known to have a power law distribution with exponent $1/2$ [see e.g. 39, p.152]. From Fig. 1 it is clear that the duration, H, of an elementary jam is $H = (D - U)/2$ and therefore retains the same power-law distribution with exponent $1/2$:

$$\Pr\{H > h\} \sim h^{-1/2} \tag{7}$$

The consequences of this are far-reaching: both the mean and variance of elementary jam durations and distance traveled diverge (to infinity on a ring road, and to values proportional to the system size on a finite road), which renders standard statistical tools inadequate. Most notably, the law of large numbers does not apply: taking a sample of jam durations to estimate mean and variance is futile because these estimates will not converge regardless of the sample size.

3.1.1. Jam sizes

Let us define a *jam* as a congestion cluster in the time-space (t, x) plane, i.e. a (t, x)-region where all vehicles are stopped. In terms of elementary jams, a jam can be characterized by a collection of, say n, of contiguous elementary jams. Turning to the discrete version of our problem for simplicity, the size S of a jam can be written as:

$$S = \sum_{i=1}^{n} H_i^*, \qquad \text{with } H_i^* \equiv H_i | H_i > H_{i-1} \tag{8}$$

where the H_i^*'s are elementary jam durations conditioned on being greater than the previous one. To see this, consider the duration H_i in Fig. 1. One can see that the duration of the elementary jam above it, H_{i-1}, which is part of the same jam, is smaller than H_i and therefore $H_1 < H_2 < \ldots < H_n$. Since the conditional distribution $\Pr\{H_i^* > h\} \sim h^{-1/2}$ is also a power law with the same exponent, $1/2$ in this case, one can expect S to also obey the same power law, regardless of the value of n [40]. We conclude:

$$\Pr\{S > s\} \sim s^{-1/2}, \tag{9}$$

which has been verified numerically; see inset in Figure 1. The experiment in the figure consistent in evolving ECA-184 for $T = 5,000$ time periods starting with an initial row of $L = 10,000$ cells, each with a probability $p = 1/2$ of being occupied. The time-space area $A = LT/2$ for cluster measurement is of triangular shape as in Figure 1, where $L = 2T$ to ensure that all points in A are able to reach the boundary. The jam sizes were plotted on a log-log plot so that the slope of the linear segment gives the exponent of the power law.

3.1.2. Exponent Near the Critical Density

To study the behavior near the critical density, we expand the experiment above to include a range of mean initial densities p. Figure 2 shows the relationship between the estimated power-law

exponent α and p in our experiment for $L = 1,000$. The downwards trend is fairly linear around the critical density, as in second-order phase transitions, and can be described approximately as

$$\alpha(p) \approx 1/2 - 4(p - 1/2) \pm 0.1, \qquad 0.45 \leq p \leq 0.55 \qquad (10)$$

The sizable but decreasing interval width of the transition, $\approx \pm 0.1$ at the critical density, is a strong indication that the exponent that best describes a particular realization is highly sensitive to the initial conditions. This makes prediction challenging, especially considering that $L = 1,000$ is a fairly large road segment of around 10 kilometers, and that the interval width increases for smaller segments, e.g. $\approx \pm 0.2$ for $L = 100$ as seen in Figure 2.
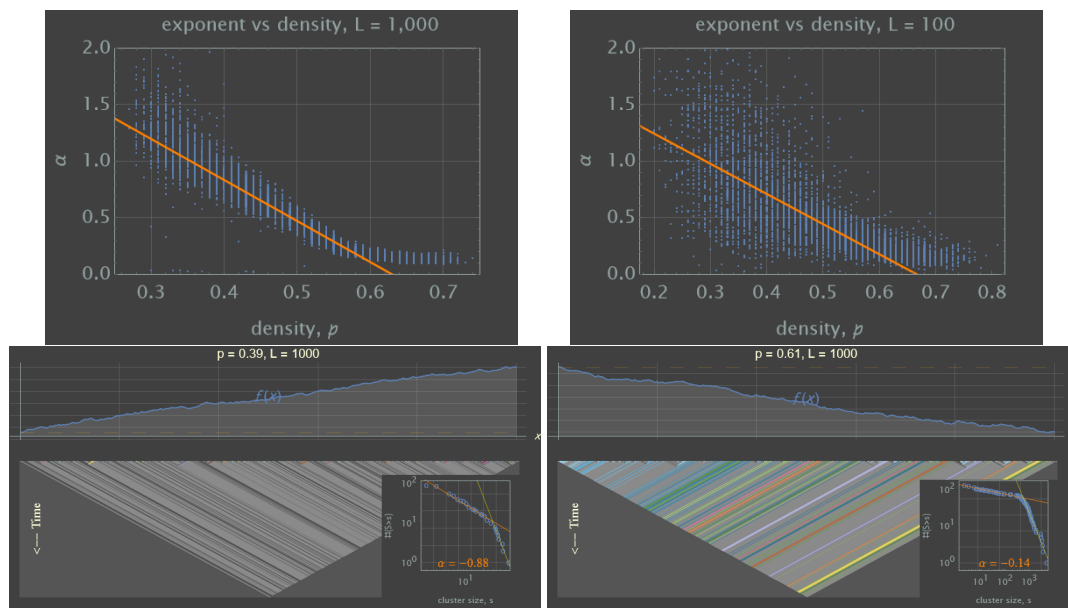
Figure 2. Top: Estimated power-law exponent α vs the initial density p for $L = 1,000$ and $L = 100$; orange regression lines $\alpha = 2.476 - 4.022p$ and $\alpha = 1.777 - 2.668p$, respectively, estimated with data in $0.45 \leq p \leq 0.55$. For lower values of the segment length results are similar but with a higher variance. Bottom: Detailed outputs for selected values of $p \neq 1/2$ for $L = 1,000$.

It is important to note that only near the critical density these power-law exponent are meaningful. Figure 2 (Bottom) shows 3 detailed outputs for selected values of p, where one can see that for $p << 1/2$ there is a clear power law relationship but with a very limited range, because in the free-flow jams are very small. For for $p >> 1/2$ the cluster distribution describes a piecewise-linear relationship, as a consequence of several clusters "percolating" to the opposite side of the triangle in the time-space plane.

4. Aggregate Measures of Performance during the Transition to Equilibrium

This section uses the results from section 3 to derive the distribution of delays, flows and speeds aggregated during the initial period where the system studied here relaxes to equilibrium. This relaxation time T was shown to be $T = L/2$ [41] for ECA 184 under the same random initial conditions in this paper, and therefore we only need to analyze the system for T time units. The time-space region for analysis in this section is taken as a triangle of area $A = TL/2 = L^2/4$ as in Figure 1, which defines all the (x, t)−points that may be affected by the initial conditions according to variational theory. In this way, our results are not restricted to a ring road and also apply to a typical roadway segment.

4.1. Delay, Flow and Speed

Let $\Psi \equiv \int_A k(x,t)\, dA$ and $\Phi \equiv \int_A q(x,t)\, dA$ be the total time traveled and total distance traveled, respectively, over a (discrete) time-space region A. By letting the symbol A also denote its area, Edie's [42] average density, flow and speed in A, k_A, q_A and v_A, respectively, are given by:

$$k_A \equiv \Psi/A, \quad q_A \equiv \Phi/A, \quad v_A \equiv q_A/k_A \tag{11}$$

These general definitions can be greatly simplified in our case because the spatiotemporal solution to our problem only involves 3 traffic states: capacity state C with (flow, density) = (1/2, 1/2), the empty or void state O with (0,0) and the jam state J with (0, 1). The total area that each state occupies in A is indicated with a subscript, and they satisfy:

$$A = A_0 + A_C + A_J, \quad A_J = A_0 \rightarrow A_C = A - 2A_J \tag{12}$$

Notice that the total delay experienced by drivers inside A is equal to A_J. This is because A_J is the only area where vehicles experience congestion and all the time inside this area is a delay (since they are stopped). The total travel time $\Psi = \int_{A_C} \frac{1}{2} \cdot dA + \int_{A_J} 1 \cdot dA = A_C/2 + A_J$, and the total distance traveled $\Phi = A_C/2$. Using (12) we have:

$$\Psi = A/2, \quad \Phi = A/2 - A_J \tag{13}$$

and therefore the total delay $A_J < A/2$ is bounded. Also,

$$k_A = 1/2, \quad q_A = 1/2 - p_J, \quad v_A = 1 - 2p_J \tag{14}$$

where $p_J = A_J/A$ correspondence to the probability that a time-space point in A is congested. Perhaps unsurprisingly, it can be seen that the total total delay A_J turns out to be the key measure of performance from which all others can be derived.

4.2. Moments and Probability Densities

The main result here is that for an area $A = L^2/4$ of triangular shape as in Figure 1, the the total delay A_J corresponds to half the area under a Brownian *excursion*, $e(x)$:

$$A_J(L) = \frac{1}{2} \int_0^L e(x)dx. \tag{15}$$

A Brownian excursion is a conditioned one-dimensional Brownian motion over the interval $(0, L)$ such that its path starts and ends at the origin and it is constrained to stay positive in between. It is important to note that the area under the excursion is also given by the area between a Brownian bridge and its minimum value in $(0, L)$. This second interpretation is precisely our context: (i) exactly at the critical density the number of empty and unoccupied cells are equal so that in our random walk $f_P(0) = f_P(L)$, which corresponds to a Brownian bridge, and (ii) from Figure 1 it can be seen that (half) the total area under the random walk and its minimum value is A_J, since by definition we have:

$$A_J = \sum_{i=1}^m H_i \tag{16}$$

where m is the total number of elementary jams in the system. Unlike (8), the correlations among the H_i's in (16) play a significant role here because the sum across all elementary jams in the system is limited by the size of the system itself, and therefore its distribution will not be a power law, as shown next.

For the unit length road segment the area under a Brownian excursion, $Z = \int_0^1 e(x)dx$ is known to have the Airy distribution [43–45]; let $g(a), 0 \leq a \leq 1$ be its PDF. Figure 3 plots the PDF of the Airy distribution, $g(a), 0 \leq a \leq 1$. Notice that it is not a heavy-tail distribution; [44] shows that its mean and standard deviation are given by $\mathrm{E}\{Z\} = \sqrt{\pi/2}/2 \approx 0.627, \mathrm{SD}\{Z\} = \sqrt{5/12 - \pi/8} \approx 0.155$. Although the Airy distribution is not exactly symmetric, one can approximate the interval where most of the realizations of Z would fall: $Z \in 0.627 \pm 0.155 = (0.472, 0.781)$.

Given the self-similarity properties of Brownian motions, we have that $A_J/L^{3/2}$ has the same distribution as $Z/2$ [43], and therefore is an invariant quantity in our problem. Most of the realizations of $A_J/L^{3/2}$ would fall in the interval $(0.236, 0.391)$. Similarly, the invariant probability of congestion is $p_J L^{1/2}$ and has the same distribution as $2Z$. Importantly, it also follows that:

$$\mathrm{E}\{A_J\} \sim L^{3/2}, \qquad \mathrm{E}\{p_J\} \sim L^{-1/2}. \tag{17}$$

To correctly interpret these insights, which are not contradictory, recall that the analysis area $A = L^2/4$ changes with the size of the segment. Therefore, doubling the size of the road increases the analysis area by a factor of 4, but the total delay A_J only by a factor of $2^{3/2} \approx 2.8$, implying economies of scale. It is not surprising then, that the probability of congestion actually decreases by a factor of $2^{-1/2} \approx 2.8/4 \approx 0.71$.

To obtain the PDF of the total delay A_J one can use the above self-similarity arguments to show that

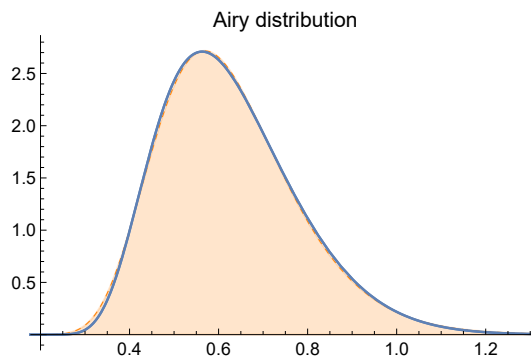$$\phi_{A_J}(a) = 2L^{-3/2}g(2aL^{-3/2}) \tag{18}$$

The distribution of all other traffic flow variables will be given by different scalings of the Airy distribution. It follows from (14) that the PDFs for these variables are given by

$$\phi_{p_J}(a) = A\phi_{A_J}(aA) = g(aL^{1/2}/2)L^{1/2}/2 \tag{19a}$$

$$\phi_{q_A}(a) = \phi_{p_J}(1/2 - a) \tag{19b}$$

$$\phi_{v_A}(a) = \phi_{p_J}((1-a)/2)/2 \tag{19c}$$

which only depend upon $L^{1/2}$.

Figure 3. PDF of the Airy distribution, $g(a), 0 \leq a \leq 1$, in blue. Although it is analytical, there is no closed-form expression for this distribution, only a complicated infinite sum [43], which is not implemented by default in most libraries/packages (the first 24 terms in this sum were considered for making this figure). The shaded area corresponds to the area under the Inverse Gaussian Distribution with mean parameter 0.625416 and shape parameter 9.82985, which can be seen to provide an accurate (and fast) approximation.

5. Discussion and Outlook

This paper has presented strong evidence that traffic flow in individual road segments exhibits all the characteristics of critical systems subject to SOC. We have seen that this critical behavior can be

traced back to the deep connection with Brownian motions formalized here thanks to the variational theory of traffic flow, establishing that even though the traffic rules are perfectly deterministic, the randomness encoded in the initial conditions is enough to generate power-law dynamics during the transient to equilibrium. Adding our willingness to maximizes travel speeds establishes SOC for the deterministic traffic flow models analyzed in this paper. We are therefore in a position now to accept Nagel and Paczuki's 1995 conjecture that traffic management strategies seeking to maximize flow are bound to cause more harm than good.

An effective control premise appears to be to *prevent* the density from reaching the critical region, e.g. by metering *well below* capacity or congestion pricing. Research is needed to devise an analytical framework able to cope with the difficulties imposed by the type of power-laws-from-fractal dynamics with exponent less than unity exposed here, which are problematic as pointed out in the introduction. The main difficulty is that traffic flow becomes "chaotic" in the sense of high sensitivity to individual initial vehicle positions, where traditional analysis tools are bound to fail. In the deterministic traffic model jams (and voids) of any duration may be expected for no particular reason near the critical density. This is where "catastrophes" take place seemingly out of nowhere: e.g., a vehicle not willing to reroute due to a spill back in front, or another one unable to change lanes in time and forced to stop, or a delivery truck suddenly double parking, all might trigger severe congestion in the whole network, in the same way earthquakes of any magnitude are expected in some countries.

Although the fractal nature of traffic flow has been known for many years, it has not been exploited in the current literature, perhaps due to the current interpretation that the basic fractal "generator" are the jams themselves, composed of smaller and smaller jams *ad infinitum* [9]. In contrast, we have seen here that the generator is a triangle similar in shape to the fundamental diagram, containing 3 traffic states: voids, capacity and jams; see Figure 1. Within each such triangle all the properties and formulas put forward in this paper apply so long as L is interpreted as the base of the triangle. Similarly, due to the symmetry property of the kinematic wave model, all the results in this paper for the jam state J transfer identically to the void state O. For example, the distribution of the duration of voids at a given location is identical to the distribution of the duration of jams at the same location; the total area of voids is also half of the area under a Brownian excursion, etc.

This paper has taken this symmetry one step further, by exploiting the connection with the theory of random walks, best illustrated using the traffic flow N-surface as in Figure A1 (right). Notably, one can see that the area under the random walk can be interpreted as the projection of this surface into the $(\Delta N, x)$-plane, and that state C is lost in this projection. It becomes clear then, that delays at the critical density can be calculated simply by integration of the initial data, without the need to solve the traffic flow model explicitly.

In the case of flow timeseries, we found compelling evidence for a Hurst exponent $H > 0.5$ (so flow fractal dimension $D = 2 - H < 3/2$) indicating long-range correlations for both cumulative counts and flow timeseries, settling the discrepancies in the existing literature between conflicting results [46] and [47] in favor of the former. It remains to be established, however, that a fractional process approach is justified for traffic flow. From Figure (A2) it is clear that the flow time series is far from a random noise symmetric around its mean value, as required by the fractional process theory. A better description of cumulative counts and flow time-series can be obtained through fractal geometry, in particular, the relationship $E\{A_J\} \sim L^{3/2}$ unveiled here. This relationship implies a fractal dimension of $3/2$ for jam areas in the time-space plane, and therefore a fractal dimension of $3/2+1=5/2$ for the N-surface, and a fractal dimension of $5/2-1=3/2$ for any slice of this surface, such as the cumulative curves or flow time-series.

We also found that SOC appears to be an intrinsic property of simple deterministic traffic flow rules, not requiring additional "cruise-control" rules as previously thought [22,23]. These references also pointed out the connection with random walks, based on numerical results that hinted to a power-law exponent close to $1/2$ for the duration of a jam. Here we have put this results on firmer grounds thanks to the connection with variational theory presented in section 3. A discrepancy found

with this earlier work is in the distribution of jam sizes (9), where they found $\alpha = 1/3$ as opposed to $1/2$ as shown here. Their result is a direct consequence of their conjecture that the number of vehicles under congestion scales with time as $t^{1/2}$, for which we found no support here.

A next step is to analyze more general initial conditions that allow for spatial correlations, such as fractional Gaussian noise processes with Hurst exponent H, where $H = 1/2$ corresponds to the Brownian white noise studied here. Another interesting direction is to expand these results for autonomous vehicles, which are not only influenced by the vehicle in front but by the 2D surrounding environment.

Finaly, the main reason why flow-maximizing strategies have not proven problematic in the traffic control literature might be explained because the traffic models typically used are meso or macroscopic, typically some numerical discretization of an ODE or PDE, such as the reservoir MFD models or variation of the cell transmission model [48]. While it is true that this model is a numerical solution method for the kinematic wave model, as any numerical scheme, it is unable to capture the nuanced sensitivity to initial conditions exposed here, which are wiped-out by (i) the mesoscopic averaging over typically long cells and/or by (ii) the considerable numerical error of these models. This argument makes the case for the use of microscopic models, but at the expense of increased computation times. Another research direction would be mesoscopic models that capture critical behavior, which appear to be nonexistent today.

Appendix A. Marginal Distribution of the N-surface
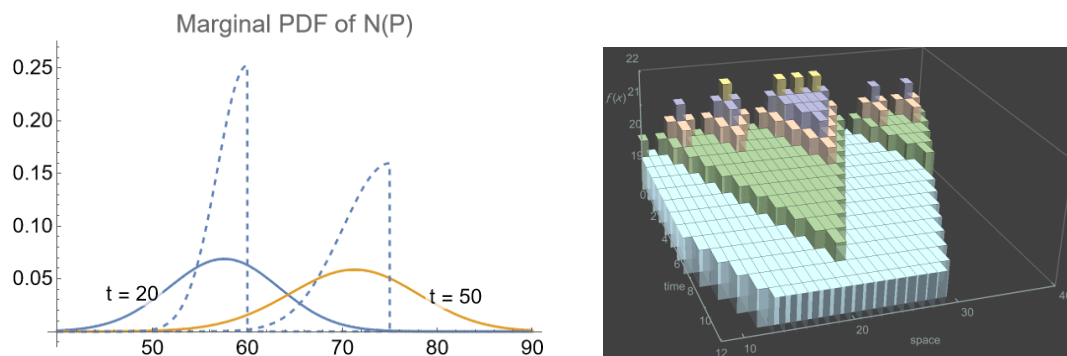
Figure A1. Left: Marginal distribution for $N(x,t)$ using $L = 200, x = L/2$ for selected times. For comparison, the half-normal distribution (A1) is included as dashed lines. Right:3D representation of the N-surface in oblique coordinates, where the third axis corresponds to deviations around the deterministic mean of the N-surface: $\Delta N(x,t) = N(x,t) - N_{\text{det}}(x,t)$, with $N_{\text{det}}(x,t) = (L - (x - t))/2$. The random walk can be seen as a slice of the deviation surface at $t = 0$, and the area under the random walk as the projection of the surface into the $(\Delta N, x)$-plane; note that state C is lost in this projection.

From the variational formulation (4) it becomes clear that the distribution of the vehicle number $N(x,t)$ is given by the minimum of the Brownian motion over the finite interval $x_U < y < x_D$. Conditional on the Brownian motion being zero at the beginning of the interval, $f_P(U) = 0$, this minimum is known to have the half-normal distribution with parameter $\sqrt{\pi/2}(x_D - x_U)$ in our case. Since $x_D - x_U = 2t$ it follows that the probability density of $N(x,t)$ conditional on $f_P(U) = n_U$ is a shifted half-normal distribution given by:

$$\phi_{N(x,t) \mid n_U}(n) = \frac{2}{\sqrt{\pi t}} \exp\left(-(n - n_U)^2/t\right), \qquad n \leq n_U \tag{A1}$$

with mean $E\{N(x,t) \mid n_U\} = n_U - \sqrt{t/\pi}$. A natural choice for n_U would be $E\{N(U)\} = (L - x + t)/2$:

$$E\{N(x,t) \mid n_U = (L - x + t)/2\} = (L - x + t)/2 - \sqrt{t/\pi} \tag{A2}$$

and one can see that the stochastic solution is below the deterministic one, i.e. $E\{N(x,t) \mid n_U\} < N_{det}(x,t)$, as expected from Jensen's inequality, but with the discrepancy $\sqrt{t/\pi}$ increasing only as $\sim \sqrt{x_D - x_U}$. The distribution (A1) is illustrated by the dashed lines in Figure A1.

The unconditional marginal distribution for $N(x,t)$ is obtained by taking expectation of (A1) with respect to the distribution of n_U, which is normal with mean $v = (L - x + t)/2$ and variance $(L - x + t)/4$:

$$\phi_{N(x,t)}(n) = (\pi(v + t))^{-1/2} \exp\left(-\frac{(n - v)^2}{v + t}\right) \text{erfc}\left((n - v)\sqrt{\frac{t}{v(v + t)}}\right) \tag{A3}$$

Unfortunately, analytical expressions for the moments of this distribution are unavailable at this time. Figure A1 (Left) illustrates the shape of this distribution at a fixed location and for different times. It can be seen that the distribution is much flatter than the half-normal distribution (A1) (dashed lines in the figure), as a consequence of large variance of n_U.

A realization of the N-surface is shown in Figure A1 (Right), which depicts a 3D representation this surface in oblique coordinates where the third axis corresponds to deviations around the deterministic mean of the N-surface: $\Delta N(x,t) = N(x,t) - N_{det}(x,t)$. The random walk can be seen as a slice of this surface at $t = 0$, and the area under the random walk as the projection of the surface into the $(\Delta N, x)$-plane; note that state C is lost in this projection.

Appendix B. Distribution of Local Flows and a Validation

A full description of the N-surface distribution, $\Pr\{N(P_1) < n_1, N(P_2) < n_2, \ldots\}$, has proven difficult to obtain, from which the distribution of flows, $q(x,t)$, could be derived. However, we have verified that the mean of the conditional (A1) and unconditional (A3) distributions are numerically very similar, and thus the mean flow can be approximated taking the time-derivative in (A2), to obtain:

$$E\{q(t)\} \approx (1 - (t\pi)^{-1/2})/2, \tag{A4}$$

where we have used $q(t)$ instead of $q(P)$ to emphasize that the flow is location-independent. Notice that this function starts at 0 at $t = 1/\pi$, increases rapidly to 0.4 by $t = 10$ and flattens out to converge asymptotically to the capacity $1/2$. (As a reference, recall that the critical headway is 2 time units.) That the flow is expected to be low at earlier times is a consequence of the system being farther from equilibrium at earlier times, which translates into an increased activity in jam interactions that produce little flow. As time goes by these interactions resolve to produce flows closer to capacity. This phenomenon can be seen in Figure 1, where it is clear that jam interactions are much more prevalent near the base of the triangle then near the bottom vertex.

Notice that the expected speed of traffic can be obtained simply by dividing (A4) by the critical density $1/2$, which coincides with the finding in [41,49] obtained by approximating the dynamics of ECA 184. This is yet another statement of universality in traffic flow, where all the models mentioned in section 1, no matter discrete or continuum, describe the same phenomena.

To gain additional insights about the distribution of flows, note from Figure 1 that the flow can be either $1/2$ in the capacity state or 0 in the void and congested states. Figure A2 shows the (oblique) cumulative counts and flow time series at 3 different locations for a typical rollout. The roughness of the cumulative curve is apparent as a result of the "spiky" variations in the flow around the capacity value $1/2$. As a first approximation to describe the stochastic processes for flows, a fractional Gaussian noise process was fitted for 2000 realizations of flows timeseries. Also, a fractional Brownian motion process was estimated for the cumulative counts. Each panel shows the estimated Hurst exponent

and the histograms summarize the results across rollouts. The average Hurst exponent for the flows $H = 0.81$ is in line with the empirical results reported for Australian motorways ($H = 0.8$, [46]) and China's highways ($H = 0.84$, [50]). This good agreement with empirical data gives a strong indication of the validity of the proposed framework.
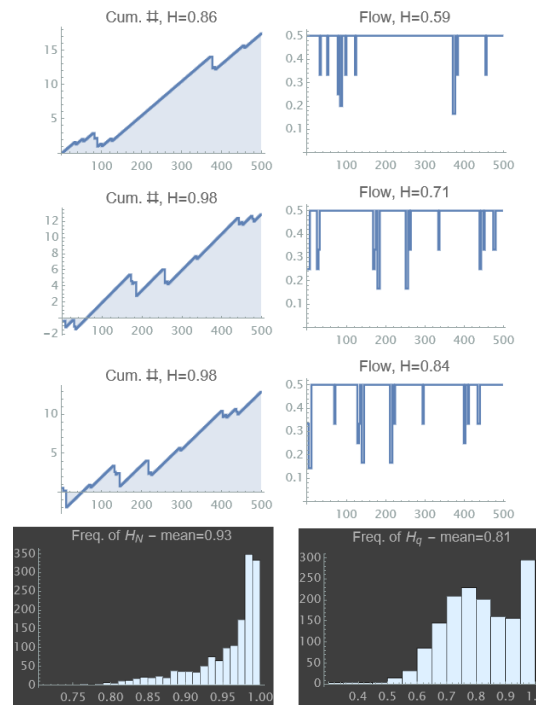
Figure A2. Oblique cumulative count $N(t) = \int_0^t (q(s) - 0.45)ds$ and flow time series at 3 different locations for a typical rollout with $L = 1000$. Notice that the flow time series have aggregated every 10 time steps for better visualization. Each panel shows the Hurst exponent estimated by fitting a fractional Brownian motion process for the cumulative counts and a fractional Gaussian noise process for the flows. The histograms summarizes the results for 2000 rollouts.

These high Hurst exponents indicate a long memory in the time series as a result of long-range spatial correlations. In this case these correlations are due to the spatial interactions brought about by the traffic model, which imposes long periods of constant capacity flow interrupted by bursts of null flow. Notice that the fractal dimension of the flow time series can be approximated by $D = 2 - H = 2 - 0.81 \approx 1.2$.

References

1. Zeng, G.; Li, D.; Guo, S.; Gao, L.; Gao, Z.; Stanley, H.E.; Havlin, S. Switch between critical percolation modes in city traffic dynamics. *Proceedings of the National Academy of Sciences* **2019**, *116*, 23–28.
2. Zhang, L.; Zeng, G.; Li, D.; Huang, H.J.; Stanley, H.E.; Havlin, S. Scale-free resilience of real traffic jams. *Proceedings of the National Academy of Sciences* **2019**, *116*, 8673–8678.
3. Zeng, G.; Gao, J.; Shekhtman, L.; Guo, S.; Lv, W.; Wu, J.; Liu, H.; Levy, O.; Li, D.; Gao, Z.; others. Multiple metastable network states in urban traffic. *Proceedings of the National Academy of Sciences* **2020**, *117*, 17528–17534.
4. Bak, P.; Tang, C.; Wiesenfeld, K. Self-organized criticality: an explanation of 1/f noise. *Phys. Rev. Lett* **1987**, *59*, 381.
5. Dănilă, B.; Harko, T.; Mocanu, G. Self-organized criticality in a two-dimensional cellular automaton model of a magnetic flux tube with background flow. *Monthly Notices of the Royal Astronomical Society* **2015**, *453*, 2982–2991.

6. Aschwanden, M.J. A macroscopic description of a generalized self-organized criticality system: Astrophysical applications. *The Astrophysical Journal* **2014**, *782*, 54.

7. Nagatani, T. The physics of traffic jams. *Reports on progress in physics* **2002**, *65*, 1331.

8. Helbing, D. Traffic and related self-driven many-particle systems. *Reviews of modern physics* **2001**, *73*, 1067.

9. Chowdhury, D.; Santen, L.; Schadschneider, A. Statistical physics of vehicular traffic and some related systems. *Physics Reports* **2000**, *329*, 199–329.

10. Nagatani, T. Traffic flow on percolation-backbone fractal. *Chaos, Solitons & Fractals* **2020**, *135*, 109771.

11. Taleb, N.N. *Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications*; STEM Academic Press, 2020.

12. Halperin, B.; Hohenberg, P. Scaling laws for dynamic critical phenomena. *Physical Review* **1969**, *177*, 952.

13. Zipf, G. Human Behaviour and the Principle of Least Effort, 1949.

14. Hurst, H.E. Long-term storage capacity of reservoirs. *Transactions of the American society of civil engineers* **1951**, *116*, 770–799.

15. Mori, T.; Smith, T.E.; Hsu, W.T. Common power laws for cities and spatial fractal structures. *Proceedings of the National Academy of Sciences* **2020**, *117*, 6469–6475.

16. West, G.B.; Brown, J.H.; Enquist, B.J. A general model for the origin of allometric scaling laws in biology. *Science* **1997**, *276*, 122–126.

17. Newman, M.E. Power laws, Pareto distributions and Zipf's law. *Contemporary physics* **2005**, *46*, 323–351.

18. Mandelbrot, B.; Hudson, R.L. *The Misbehavior of Markets: A fractal view of financial turbulence*; Basic books, 2007.

19. Kobayashi, Y.; Takayasu, H.; Havlin, S.; Takayasu, M. Robust Characterization of Multidimensional Scaling Relations between Size Measures for Business Firms. *Entropy* **2021**, *23*, 168.

20. Watanabe, H.; Takayasu, H.; Takayasu, M. Relations between allometric scalings and fluctuations in complex systems: The case of Japanese firms. *Physica A: Statistical Mechanics and its Applications* **2013**, *392*, 741–756.

21. Clauset, A.; Shalizi, C.R.; Newman, M.E. Power-law distributions in empirical data. *SIAM review* **2009**, *51*, 661–703.

22. Nagel, K.; Paczuski, M. Emergent traffic jams. *Physical Review E* **1995**, *51*, 2909.

23. Paczuski, M.; Nagel, K. Self-organized criticality and 1/f noise in traffic. Traffic and granular flow. World Scientific Singapore, 1996, p. 73.

24. Nagel, K.; Schreckenberg, M. A cellular automaton model for freeway traffic. *Journal de physique I* **1992**, *2*, 2221–2229.

25. Daganzo, C.F. A Variational Formulation of Kinematic Wave Theory: basic theory and complex boundary conditions. *Transportation Research Part B* **2005**, *39*, 187–196.

26. Nagel, K. Personal communication, 2021.

27. Schadschneider, A. Personal communication, 2021.

28. Nagel, K.; Rasmussen, S.; Barrett, C.L. Network traffic as a self-organized critical phenomena. Technical report, Los Alamos National Lab., NM (United States), 1996.

29. Rieser, M.; Nagel, K. Network breakdown "at the edge of chaos" in multi-agent traffic simulations. *The European Physical Journal B* **2008**, *63*, 321–327.

30. Olmos, L.E.; Çolak, S.; Shafiei, S.; Saberi, M.; González, M.C. Macroscopic dynamics and the collapse of urban traffic. *Proceedings of the National Academy of Sciences* **2018**, *115*, 12654–12661.

31. Lighthill, M.J.; Whitham, G.B. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1955**, *229*, 317–345.

32. Richards, P.I. Shock waves on the highway. *Operations research* **1956**, *4*, 42–51.

33. Newell, G.F. A Simplified Car-Following Theory : a Lower Order Model. *Transportation Research Part B* **2002**, *36*, 195–205.

34. Daganzo, C.F. In Traffic Flow, Cellular Automata = Kinematic Waves. *Transportation Research Part B* **2006**, *40*, 396–403.

35. Laval, J.A.; Chilukuri, B.R. Symmetries in the kinematic wave model and a parameter-free representation of traffic flow. *Transportation Research Part B: Methodological* **2016**, *89*, 168 – 177.

36. Wolfram, S. Cellular automata as models of complexity. *Nature* **1984**, *311*, 419.

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

38. Hopf, E. On the Right Weak Solution of the Cauchy Problem for a Quasilinear Equation of First Order. *Indiana Univ. Math. J.* **1970**, *19*, 483–487.

39. Schroeder, M. *Fractals, chaos, power laws: Minutes from an infinite paradise*; Courier Corporation, 2009.

40. Zaliapin, I.; Kagan, Y.Y.; Schoenberg, F.P. Approximating the distribution of Pareto sums. *Pure and Applied geophysics* **2005**, *162*, 1187–1228.

41. Fuks, H. Solution of the density classification problem with two cellular automata rules. *Physical Review E* **1997**, *55*, R2081.

42. Edie, L.C. Discussion of Traffic Stream Measurements and Definitions. 2nd Int. Symp. on Transportation and Traffic Theory; , 1965; pp. 139–154.

43. Majumdar, S.N.; Comtet, A. Airy distribution function: from the area under a Brownian excursion to the maximal height of fluctuating interfaces. *Journal of Statistical Physics* **2005**, *119*, 777–826.

44. Janson, S. Brownian excursion area, Wright's constants in graph enumeration, and other Brownian areas. *Probability Surveys* **2007**, *4*, 80–145.

45. Agranov, T.; Zilber, P.; Smith, N.R.; Admon, T.; Roichman, Y.; Meerson, B. Airy distribution: Experiment, large deviations, and additional statistics. *Physical Review Research* **2020**, *2*, 013174.

46. Chand, S.; Aouad, G.; Dixit, V.V. Long-Range Dependence of Traffic Flow and Speed of a Motorway: Dynamics and Correlation with Historical Incidents. *Transportation Research Record* **2017**, *2616*, 49–57.

47. Krause, S.M.; Habel, L.; Guhr, T.; Schreckenberg, M. The importance of antipersistence for traffic jams. *EPL (Europhysics Letters)* **2017**, *118*, 38005.

48. Daganzo, C.F. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B* **1994**, *28*, 269–287.

49. Fuks, H.; Boccara, N. Generalized deterministic traffic rules. *International journal of modern physics C* **1998**, *9*, 1–12.

50. Shang, P.; Wan, M.; Kama, S. Fractal nature of highway traffic data. *Computers & Mathematics with Applications* **2007**, *54*, 107–116.