Buses of Cuernavaca - an agent-based model for universal random matrix behavior minimizing mutual information

Piotr Warchoł

M. Smoluchowski Institute of Physics, Jagiellonian University, PL–30–348 Cracow, Poland
(Dated: October 2, 2017)

The public transportation system of Cuernavaca, Mexico, exhibits random matrix theory statistics [1]. In particular, the fluctuation of times between the arrival of buses on a given bus stop, follows the Wigner surmise for the Gaussian Unitary Ensemble. To model this, we propose an agent-based approach in which each bus driver tries to optimize his arrival time to the next stop with respect to an estimated arrival time of his predecessor. We choose a particular form of the associated utility function and recover the appropriate distribution in numerical experiments for a certain value of the only parameter of the model. We then investigate whether this value of the parameter is otherwise distinguished within an information theoretic approach and give numerical evidence that indeed it is associated with a minimum of averaged pairwise mutual information.

PACS numbers: 02.10.Yn, 02.50.Ey, 89.70.+c , 05.90.+m

I. INTRODUCTION

As a galvanizing example of a socio-economic complex system exhibiting random matrix statistics, the buses of Cuernavaca [1] are a poster child for universality [2]. Thus, their behavior, like other phenomena of its kind, is not only interesting for its own sake, as a realization of some physical dynamical system, but it is also intriguing due to its connection to random matrix theory. Moreover, after 17 years since the publishing of the seminal paper, the associated phenomenon is still being rediscovered in different systems like cars on a motorway [3] and subway trains in New York city [4].

As described by Krbálek and Šeba, in Cuernavaca, the buses do not have a given schedule. In contrast to most other cities, their drivers don’t work for a single company. Instead, they operate according to their own interest, optimizing the amount of money earned while following a given route. Therefore, they strive to serve as many passengers as possible. This on the other hand means, that they tend not to arrive on a given bus stop immediately after their predecessors. To achieve that, they learn the information about the time of departure of the previous bus from collaborators stationed at the bus stop they arrive at. This mechanism seems to lead to the recreation of random matrix statistics for the unfolded time intervals between arrivals of consecutive buses to a given bus stop, and the number variance (the variance of the number of buses arriving to a bus stop in a given time interval).

Many models have been proposed to describe this phenomenon. In particular, the authors of [6], famously formulate the problem as a set of independent, rate one, Poisson processes conditioned not to intersect. A cellular automaton and a novel matrix model, the Damped Unitary Ensemble are proposed in [7] and [8] respectively. In the first ([6]) case the relevant GUE statistics are recovered through an analytical calculation, yet the explanatory power of the model is diminished by the complicated mathematics involved. The two other models, on the other hand are dependent on a parameter, which drives the statistics from Poissonian to random matrix (GUE) like behavior. In the case of the cellular automaton, it is a ratio of the number of bus stops at which the bus drivers receive information and the number of all the bus stops. The size of the deterministic, off-diagonal elements of the Damped Unitary Ensemble matrices plays that role in the second case.

We see the same transition in the aforementioned case of highway cars, with the traffic density as the driving factor, and for the New York city subway, where an effective value of Coulomb potential is assigned to the different behaviors.

Inspired by an approach of combining Game Theory and Random Matrix Theory to tackle performance optimization of multi-user telecommunication systems [9], we introduce a novel model for simulating the bus system of Cuernavaca. In it, at each bus stop, each of the drivers optimizes his expected arrival time to the next stop with the use of a simple utility function. This function has one parameter which plays exactly the same role as the ones described above. Namely, it drives the statistics of the bus trajectories from Poissonian, through GUE to Gaussian. We then go one step further, and ask, whether the value of this parameter, which effectively measures the strength of the repulsive potential and for which we obtain the Random Matrix behavior, is otherwise singled out. To answer this question, we turn to information theory. It turns out that this particular value is associated with a minimum of averaged mutual information between the trajectories of neighboring buses (Averaged Pairwise Mutual Information - APMI), measured at a given bus stop. We conjecture that it is this feature of the system that makes the Cuernavaca buses exhibit GUE statistics.

This paper is organized as follows. We start by characterizing the proposed model and report the results of the simulation for the nearest neighbor spacing distribution and the number variance. This is followed by introducing the pairwise mutual information for this context and a presentation of the results of its estimation. We finish the paper with a brief conclusion and some conjectures.
II. THE MODEL

First, let us describe our model for the bus system of Cuernavaca. Consider \( N \) bus drivers traveling along a circular route with \( K \) ordered stops. At each stop \( k - 1 \), each of the drivers \( n \), in turn learns their time of arrival \( t_{n,k-1} \) and the time of arrival of their predecessor \( t_{n-1,k-1} \). Based on that, and the knowledge of the route (namely the average time to travel between stops \( k - 1 \) and \( k \) denoted as \( t_k \)) , they assess at what time \( t_{n,k} \) to arrive at the next bus stop to maximize their earnings. Thus, each maximizes some utility function \( u_{n,k} \) for the arrival of the \( n \)'th bus to the \( k \)'th stop, with respect to the intended arrival time \( t_{n,k} \). The amount of people entering the bus and the arrival time itself (due to unexpected occurrences on the route) will in the end be random variables. We will however neglect that and work with the utility function not its expectation. Thus, in the end, at each bus stop, each bus driver solves

\[
\frac{\partial u_{n,k}}{\partial t_{n,k}} = 0 \tag{1}
\]

to find the optimal arrival time.

Now, let us design the utility function. We know that the bus drivers will strongly avoid not taking anybody at the stop, which happens when the time distances \( (t_{n,k} - t_{n-1,k}) \) between arrivals are very short. On the other hand, if they take a sufficiently large number of people, they won’t care as much about taking a few more. This (as well as Random Matrix Theory) suggests taking the logarithmic function to represent such behavior. It will be the repulsive core of the potential. The difference between the arrival times will not be expressed in terms of \( t_{n-1,k} \) as this is not known for the \( n \)'th driver yet (at bus stop \( k - 1 \)). Instead, it is expressed with the help of the average travel time between the stops and the known arrival time of bus \( n - 1 \) to the stop \( k - 1 \) as \( t_{n,k} - t_{n-1,k-1} - t_k \). The journey, moreover, cannot be too long and the passengers value how predictable is the time it is going to take. We will use the harmonic potential to represent the fact that the travel time should be close to the average. Thus, we set:

\[
u_{n,k} = b \ln (t_{n,k} - t_{n-1,k-1} - t_k) - (t_{n,k} - t_{n-1,k-1} - t_k)^2 \tag{2}\]

and therefore each of the drivers, solves the following quadratic equation

\[
b = 2(t_{n,k} - t_{n-1,k-1} - t_k)(t_{n,k} - t_{n-1,k-1} - t_k) \tag{3}\]

and chooses the solution consistent with \( t_{n,k} > t_{n-1,k} \). This can be seen as \( N \), mutually dependent nonlinear maps. \( b \) is some yet unknown parameter, which we assume to be the same for all the bus drivers in the model and which marks the strength of the logarithmic repulsion with respect to the harmonic potential. The presented approach is convenient as we can explicitly solve the quadratic equation and therefore, as we describe in the next section, simulating this model on a computer is straightforward.

Thus, there are many agents (bus drivers), each optimizing the behavior with respect to the agent in front. Moreover, a change in the arrival time with respect to the optimal one, for some bus driver, will by definition result on average with a smaller pay-off. The realization of this scheme is thus a form of Nash equilibrium [11] - each player, at every bus stop can choose from a continuous set of strategies associated with the travel times. Note, that this does not generalize to a Wardrop equilibrium [11, 12] as the number of drivers grow, because the choice of the strategy of one, directly influences the situation of only the one immediately behind him. Finally, the optimization is done individually and for each bus stop separately, hence there might be a global equilibrium possible, for which the total earnings of all the bus drivers are higher then the total earning resulting with using this optimization method (this being a hallmark of a decentralized systems).

III. RESULTS OF THE SIMULATION

Now, our goal is to compare this model to the real bus system of Cuernavaca. To this end, for the simulated system we will compute the two statistical properties studied in the original paper by Krba{é}ek and Seba - the unfolded nearest neighbor spacing distribution and the so called number variance. Under the condition that, on average, the time distances between the buses arriving at a given stop are one (realized by the unfolding procedure), the former is the distribution \( P(s) \) of time distances \( s \) between the arrivals of the neighboring buses. The latter \( \langle N(t) \rangle \), under the same condition, measures the variance of the number of buses arriving in a time interval of length \( t \). For random matrices of the Gaussian Unitary Ensemble, these are given by the Wigner surmise approximation:

\[
P(s) = \frac{32}{\pi^2} s^2 e^{-\frac{2}{\pi} s^2} \tag{4}\]

and

\[
\langle N(t) \rangle \approx \frac{1}{\pi^2} (\log 2\pi t + \gamma + 1) \tag{5}\]

respectively [13], and where \( \gamma \) is the Euler constant.

FIG. 1. Eleven sample bus trajectories, after unfolding, for three different values of \( b \). The horizontal axes represents the consecutive bus stops, and the vertical one, the unfolded arrival times.

Thus, for our simulation, we generate 100 buses traveling around a cyclic root with 100 stops distributed so that the
The travel time between the adjacent stops ($\tau_k$) is given by a uniform distribution, here on the interval $[1, 1.5]$. The latter sets the time scale of the experiment. We initialize the simulation by setting the times of arrival of the first bus (labeled with 1) to all the stops on the first cycle. These are normally distributed, with mean $t_{1,k-1} + \tau_k$ and variance 0.05, which constitutes 4% of the average $\tau_2$ across the bus stops. $t_{1,1}$ is set to be 0. Then, we set the times of arrival to the first stop in the first cycle for each of the rest of the bus drivers. These start their journeys in equal time intervals such that on average, the last leaves the first stop when the first driver arrives (on average) at the last one. The rest of the simulation is performed through the optimization method described above and with a cyclic boundary condition, so that when the bus driver leaves bus stop $K$ it goes towards the one labeled by 1. The arrival time is at each step perturbed, with respect to the one resulting from the optimization procedure, by a random number form the normal distribution with mean 0 and variance $\sigma = 0.05$ (again, 4% of the average $\tau_2$ across the bus stops). Here, we keep $\sigma$ fixed, however this is a parameter that has some amount of qualitative influence on the results - we will come back to this later. For now we just mention that it has to be small, so that bus overtakings due to noise are extremely rare.

![FIG. 2. Nearest neighbor spacing distribution. The red line is the analytic approximation (the so-called Wigner surmise) of the GUE random matrix result. The blue crosses are results obtained in [1] for the Cuernavaca buses whereas the blue histogram is the results of the numerical simulations for $b = 0.018$.](image)

Altogether, we perform 100 such experiments, with 32 cycles, for each given value of $b$ (between 0.0002 and 0.1602 with a step of 0.002). The distribution of arrival times for a given stop in a given cycle is already, on average, uniform. Thus, what the unfolding procedure effectively does, it rescales the arrival times so that they are all, on average, uniformly distributed on an interval from 0 to $N$.

We showcase some of the unfolded bus trajectories generated with the numerical simulations in figure 2. For large values of $b$, up to some small fluctuations, the buses travel in unison, whereas for small ones, there is little coordination on larger timescales. The associated nearest neighbor distance distributions can be seen in figures 2 and 3. The former plot depicts it for the value of $b$ for which the ($\beta = 2$) Wigner surmise and the histogram match best. The skewness of the distribution obtained with the model is the result of the value of $\sigma$ being small but finite. Making it smaller reduces the skewness to the point of a very good match with the random matrix result. This would also mean a rescaling for $b$ (the essence of the information measure result below does not change however) [14]. As can bee seen in the second plot, for significantly smaller $b$’s, the distribution approaches the Poissonian statistics, as the harmonic potential dominates. For large values of $b$ it becomes Gaussian. Then, the logarithmic part of the ‘potential’ has the upper hand and this behavior comes from the random (normal) perturbation of the arrival times with respect to the optimal arrival time the bus drivers strived to achieve.

![FIG. 3. Nearest neighbor spacing distribution. The red line is the analytic approximation (the so-called Wigner surmise) of the GUE random matrix result. The histograms are the results of the numerical simulations for $b = 0.0002$ (green) and $b = 0.1602$ (purple). The dashed green and purple lines mark the exponential and normal (best fit, with mean 1) distributions respectively.](image)

The equivalent of the random matrix theory number variance is the variance of the distribution of the number of buses arriving at a certain bus stop in a given time interval. This is
the neighboring buses, in terms of entropy is:

\[
I(W_n : W_{n+1}; b) = H(W_n; b) + H(W_{n+1}; b) - H(W_n, W_{n+1}; b).
\]

Finally, the averaging is done over all the vehicles. Therefore, APMI measures the reduction of uncertainty in the arrival time (subject to the unfolding procedure done across the arrival times to a given bus stop) of one of the two neighboring buses which stems from the knowledge of the arrival time (to the same bus stop) of the other and vice-versa. In other words, it is the information shared by these two random variables.

To calculate it, we use the unfolded arrival times of the last 29 (of the generated 32, to allow for ‘thermalization’) cycles of 100 stops, for each nearest neighbor pairs of the 100 buses and perform this experiment 100 times. The Kraskov-Stoegbauer-Grassberger estimation method [18], as implemented in the JIDT package [19] as Kraskov algorithm 1 is utilized.

![Figure 5](image)

FIG. 5. In red, the averaged pairwise mutual information (measured in nats) for values of \( b \) in 0.002 intervals between 0.0002 and 0.1602. We plot the error bars representing the standard error of the mean of the results from all the experiments. The vertical blue line marks \( b = 0.018 \).

We can see the results in figure[5]. The pairwise mutual information, within the studied spectrum of \( b \), reaches its minimum close to the value of \( b \) related to the GUE. Thus, as advertised earlier, we claim that the bus drivers of Cuernavaca, while maximizing their payoff, they minimize the mutual information with respect to the neighboring buses. This in turn leads to the manifestation of the GUE statistics of the nearest neighbor spacing distribution.

Let us note that it is clear why APMI is large for large values of \( b \) - the bus trajectories are rigid with respect to each other and the amount of shared information is high. Why is the value large for very small \( b \)’s? This is because then, the overall entropy of the trajectories is large. In fact the relative APMI (APMI divided by the averaged entropy of the particular trajectories, namely the fraction of shared information and the uncertainty) is monotonic across the spectrum of \( b \), rising with growing \( b \), in accordance with intuition.
V. SUMMARY AND CONCLUSIONS

In this paper we have proposed a straightforward, agent-based approach for modeling the bus system of Cuernavaca. The resulting nearest neighbor distance distribution and number variance compare favorably to the associated results for random matrices and the bus system itself. The proposed model has one adjustable parameter, however we showed that choosing its value is linked to a minimization of averaged pairwise mutual information. This enables us to link the global statistical behavior of this complex system to, both, a maximization of some profit function of individual drivers and a minimization of the information shared between the nearest neighbors. Note however, that the mechanism behind the latter is still unclear. Perhaps there is some optimization process not taken into account, otherwise one is tempted to see the minimization of the APMI as a version of a kind of Maximum Entropy Principle at play for complex interacting systems. Either way, it would be very interesting if the same effect is visible in the parameter dependent models mentioned in the introduction.

Finally, we conjecture, that the studied pairwise mutual information will also have a minimum for stochastic processes generating arbitrary-β ensembles, as a function of β, for β = 2. Perhaps the constraint of Gaussianity can be also alleviated. If this is true, it may have important consequences for systems which exhibit local random matrix and semi-Poisson statistics and may help with understanding of the universality properties of the Gaussian Unitary Ensemble. We plan to verify this with numerical experiments.

VI. ACKNOWLEDGMENTS

We would like to acknowledge the funding of the Polish National Science Centre through the projects DEC-2011/02/A/ST1/00119 and 2016/21/D/ST2/01142. Various discussions on the subject with Maciej A. Nowak, Jacek Grela and Jeremi Ochab are greatly appreciated.

[1] M. ˇKrbalek and P. ˇSeba, "The statistical properties of the city transport in Cuernavaca (Mexico) and random matrix ensembles", J. Phys. A 33 (2000) 229234.
[2] F. Dyson, foreword to The Oxford Handbook of Random Matrix Theory edit by G. Akemann, J. Baik, P. Di Francesco (Oxford University Press, 2011).
[3] A. Y. Abul-Magd "Modeling highway-traffic headway distributions using superstatistics", Phys. Rev. E 76 (2007) 057101.
[4] A. Jagannath and T. Trogdon, "Random matrices and the New York City subway system", arxiv:1703.02537.
[5] Through the unfolding procedure the mean distance between adjacent variable in the sample becomes 1.
[6] J. Baik, A. Borodin, P. Deift and T. Suidan, "A model for the bus system in Cuernavaca (Mexico)" J. Phys. A 39 (2006) 8965.
[7] M. ˇKrbalek and P. ˇSeba, "Headway statistics of public transport in Mexican cities", J. Phys. A 36 (1) (2002) L7.
[8] M. ˇKrbalek and T. Hobza, "Inner structure of vehicular ensembles and random matrix theory", J. Phys. A 380 (1) (2016) 1839.
[9] N. Bonneau, M. Debbah, E. Altmana, T. Boulognea, R. El-Azouzia, T. Jiménezb and L. Wynterc, "A survey on networking games in telecommunications", Computers and Operations Research 33 (2006) 286.
[10] We decided to keep σ at 4% relative to the average time distance between the buses as this seems realistic and because we wanted to showcase the skewness effect.
[11] E. Altmana, T. Boulognea, R. El-Azouzia, T. Jiménezb and L. Wynterc, "A survey on networking games in telecommunications", Computers and Operations Research 33 (2006) 286.
[12] J. G. Wardrop, "Some theoretical aspects of road traffic research communication networks" Proceedings of the Institution of Civil Engineers, Part 2 1 (1952) 325.
[13] M. L. Mehta, Random Matrices (Academic Press, New York, 1991), 2nd ed.
[14] A. Kraskov, H. Stadeler, E. Altmana, T. Boulognea, R. El-Azouzia, T. Jiménezb and L. Wynterc, "Information Flow in a Kinetic Ising Model Peaks in the Disordered Phase", Phys. Rev. Lett. 111 (2013) 177203.
[15] A. Kraskov, H. Stadeler, E. Altmana, T. Boulognea, R. El-Azouzia, T. Jiménezb and L. Wynterc, "Information Flow in a Kinetic Ising Model Peaks in the Disordered Phase", Phys. Rev. Lett. 111 (2013) 177203.
[16] A. Kraskov, H. Stadeler, E. Altmana, T. Boulognea, R. El-Azouzia, T. Jiménezb and L. Wynterc, "Information Flow in a Kinetic Ising Model Peaks in the Disordered Phase", Phys. Rev. Lett. 111 (2013) 177203.
[17] A. Kraskov, H. Stadeler, E. Altmana, T. Boulognea, R. El-Azouzia, T. Jiménezb and L. Wynterc, "Information Flow in a Kinetic Ising Model Peaks in the Disordered Phase", Phys. Rev. Lett. 111 (2013) 177203.