

PAPER • OPEN ACCESS

# Entropy production of selfish drivers: implications for efficiency and predictability of movements in a city

To cite this article: Indaco Biazzo *et al* 2021 *J. Phys. Complex.* **2** 035026

View the [article online](#) for updates and enhancements.

## You may also like

- [Non-thermal transitions in a model inspired by moral decisions](#)  
Roberto C Alamiño
- [Evaluating the impact of selfish behaviors on epidemic forwarding in mobile social networks](#)  
Yahui Wu, Su Deng and Hongbin Huang
- [Hop limited epidemic-like information spreading in mobile social networks with selfish nodes](#)  
Yahui Wu, Su Deng and Hongbin Huang

## OPEN ACCESS



CrossMark

## RECEIVED

15 May 2021

## REVISED

1 September 2021

## ACCEPTED FOR PUBLICATION

14 September 2021

## PUBLISHED

30 September 2021

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



## PAPER

## Entropy production of selfish drivers: implications for efficiency and predictability of movements in a city

Indaco Biazzo<sup>1</sup>, Mohsen Ghasemi Nezhadhighi<sup>2</sup> and Abolfazl Ramezanpour<sup>2,3,\*</sup> <sup>1</sup> Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy<sup>2</sup> Department of Physics, School of Sciences, Shiraz University, Shiraz 71454, Iran<sup>3</sup> Leiden Academic Centre for Drug Research, Faculty of Mathematics and Natural Sciences, Leiden University, PO Box 9500-2300 RA Leiden, The Netherlands

\* Author to whom any correspondence should be addressed.

E-mail: [aramazanpour@gmail.com](mailto:aramazanpour@gmail.com)**Keywords:** transport efficiency, entropy production, predictability**Abstract**

Characterizing the efficiency of movements is important for a better management of the cities. More specifically, the connection between the efficiency and uncertainty (entropy) production of a transport process is not established yet. In this study, we consider the movements of selfish drivers from their homes (origins) to work places (destinations) to see how interactions and randomness in the movements affect a measure of efficiency and entropy production (uncertainty in the destination time intervals) in this process. We employ realistic models of population distributions and mobility laws to simulate the movement process, where interactions are modelled by dependence of the local travel times on the local flows. We observe that some level of information (the travel times) sharing enhances a measure of predictability in the process without any coordination. Moreover, the larger cities display smaller efficiencies, for the same model parameters and population density, which limits the size of an efficient city. We find that entropy production is a good order parameter to distinguish the low- and high-congestion phases. In the former phase, the entropy production grows monotonically with the probability of random moves, whereas it displays a minimum in the congested phase; that is randomness in the movements can reduce the uncertainty in the destination time intervals. The findings highlight the role of entropy production in the study of efficiency and predictability of similar processes in a complex system like the city.

**1. Introduction**

Human mobility has changed completely over the past two centuries, starting in Britain, in the early nineteenth century, from the dramatic expansion of roads and railways networks [1]. Studies aiming at describing and understanding the human mobility in space and time started some decades after that [2], and more quantitative and theoretical approaches appear starting from the 1940s [3–5]. In recent years, the great improvements of computational resources and data gathering due to the ICT revolution has given new impulses to these studies [6, 7].

The widespread adoption of mobile phones in early 2000s and subsequently of smartphones increased enormously the amount of data available about movements in city. These data fostered the appearance of new researches describing and characterizing urban displacement [8–12]. Private vehicles, until today, cover the majority of trips in large parts of the world [13]. There are however multiple drawbacks associated with this situation. Due to the increased private demands of displacements, gridlocks are rising every year [14–16], leading to large economic loss [14] and increasing urban air pollution [17]. In this regards, large investment in transport infrastructure are mandatory, but also better strategies to understanding and modelling of such complex systems can help to mitigate the undesirable effects. Agent based models have been developed in order to capture the main aspects of the phenomenon, ranging from simple models (see references in [7]) to very complicated and computationally demanding ones [18, 19]. Recently, the pervasive smartphone adoption and

the consequent use of navigation apps have changed the movement habit of individuals, and not always for the better [20].

In this work we want to see how interactions of agents and associated entropy productions influence predictability and efficiency of movements in city. We study a dynamical system with agents moving according to a selfish routing algorithm [21, 22] on a two-dimensional square lattice, with some reasonable models of population and destinations distributions which can reproduce the essential features of the empirical data [11, 23]. There is no cooperation among the agents and by selfish we just mean that the agents tend to take the expected shortest-time path according to the available information. We consider two kinds of interactions among the agents: spatial and temporal interactions. The former interactions affect the travel times of street segments which is correlated with the flows of agents. In the latter interactions, the travel information of previous days are taken into account in the selfish routing algorithm.

We shall see how the predictability and efficiency of the movements depend on the strength of interactions. Two measures of predictability are considered: the entropy production and a distance between distributions of the expected and actual travel times. The measure of efficiency was introduced in [21] as the inverse of the travel time per the total number of trips. In this study, however, we simulate the movements of many interacting agents, which can also deviate from their optimal trajectories randomly at every time step. We observe that in general the efficiency and predictability of the process are diminished by increasing the interactions in space and time. In principle, sharing the travel information with selfish drivers in the absence of any coordination could lead to traffic ‘chaos’ [20]. The phenomenon in general is known as the tragedy of commons or the price of anarchy in a population of selfish agents [24–26]. We find that very small but nonzero levels of temporal interactions can in fact enhance the predictability by reducing the differences between the expected and actual travel times. Moreover, for large values of spatial interactions, a bit of randomness in the movements results in smaller entropy production by avoiding the congested lines. For the same reason, the efficiency displays a maximum at small values of randomness when the spatial and temporal interactions are sufficiently strong.

The paper is structured as follows. In section 2, we introduce the models and define the main quantities of the study. In section 3, we report the results of numerical simulations and discuss the findings and their consequences. The concluding remarks are given in section 4.

## 2. Models and settings

In this section, we present the main definitions and methods which are used to model the network flow dynamics. Consider a city of  $N$  sites with local populations  $\{m_a : a = 1, \dots, N\}$  and total population  $M = \sum_a m_a$ . The connectivity graph of the city is given by  $G(V, E)$  where  $V$  is the set of sites and  $E$  is the set of directed edges  $(ab)$ . Here we take a two dimensional square lattice of size  $N = L \times L$ , where all edges have the same length and the same free travel times. We use the model introduced in reference [11] to produce reasonable population distributions for the model cities. This is a preferential growth model which starts with a seed of population at the center of the lattice. At each step, a site is chosen with probability  $p_a \propto (m_a + c_0)$  and a unit of population is added only if there exists a populated site at distance less than  $l_0$ . In the following, we take  $c_0 = 1$ ,  $l_0 = 1$ , and the population density is fixed to  $M/N = 10^3$ .

Given the population distribution  $m_a$ , we use the following mobility law [23] to construct the flux of movements  $m_{a \rightarrow b}$  from origins  $a$  to destinations  $b$ ,

$$m_{a \rightarrow b} = m_a p_{a \rightarrow b} = m_a \frac{m_b / M(r_{ab})}{\sum_{c \neq a} m_c / M(r_{ac})}, \quad (1)$$

where  $M(r_{ab})$  is the population in the circle of radius  $r_{ab}$  centred at site  $b$ . The probability  $p_{a \rightarrow b}$  of going from  $a$  to  $b$  is proportional to  $m_b / M(r_{ab})$ , which can be interpreted as the attractiveness of site  $b$  for an individual at site  $a$ .

Finally, the flows of movements on edges  $(ab) \in E$  are determined by a flux distribution dynamics, as follows.

### 2.1. The movement process

Here we describe the process of moving from the origins (Os) to destinations (Ds) in a single day:

- The starting times of the OD trips are distributed uniformly in the origin time interval  $\Delta T_O$ . We assume that  $\Delta T_O$  is the same for all origins. In each time step, the time increases by  $\Delta t = 1$ . Driver  $i$  starts its trip and becomes active at time  $t_O(i) \in \Delta T_O$ . The trip will become inactive when the driver reaches its destination at time  $t_D(i) \in \Delta T_D$ . The destination time interval  $\Delta T_D$  is determined by the system structure and dynamics (see figure 1). The arrival times of the drivers at destination site  $a$  determine the

destination time interval of that site  $\Delta T_D(a)$ . The travel time from origin to destination for driver  $i$  is denoted by  $\Delta t_{OD}(i) = t_D(i) - t_O(i)$ .

- An active driver  $i$  at site  $a$  chooses the next site as follows: with probability  $\alpha$  the next site is chosen randomly and uniformly from the set of neighbouring sites. With probability  $1 - \alpha$  the neighbour that minimizes the expected travel time to the destination  $D(i)$  is selected. Here, the probability  $\alpha$  controls the degree of selfishness of individuals. The expected travel time on each directed edge  $(ab)$  in day  $d$  is denoted by  $\tilde{t}_{ab}(d)$ . The expected travel times in day  $d$  are estimated by using the actual travel times  $t_{ab}(d - 1)$  from the previous day:

$$\tilde{t}_{ab}(d) = \lambda t_{ab}(d - 1) + (1 - \lambda) \tilde{t}_{ab}(d - 1). \quad (2)$$

For the initial day  $\tilde{t}_{ab}(0) = t_{ab}(0)$ , where the  $t_{ab}(0)$  are the travel times for free lines (with no flow). We shall assume that all edges in the two-dimensional lattice are the same and set  $t_{ab}(0) = 1$ . In practice  $t_{ab}(0)$  is given by the length of the line and the velocity limit. Note that when  $\lambda = 0$  the expected travel times do not change with day and the drivers always choose the shortest path according to the  $t_{ab}(0)$ .

- Let flow  $F_{ab}(t)$  be the number of drivers that enter edge  $(ab)$  at time step  $t$ . Given the flows, the actual travel times are obtained from

$$t_{ab}(F_{ab}) = t_{ab}(0) \left( 1 + g \left( \frac{F_{ab}}{F_{ab}^*} \right)^\mu \right), \quad (3)$$

to model the influence of flows on the travel times [27–29]. There are other models which try to describe the relation between the observed travel times and the flows [30]. The above relation allows us to control the strength of interactions by the parameter  $g$  and the degree of nonlinearity in the interactions by the  $\mu$ . The actual travel time  $t_{ab}$  is the time that a person spends on edge  $(ab)$ . Here  $F_{ab}^*$  is a measure of the line capacity. For simplicity, we assume that  $t_{ba}(0) = t_{ab}(0)$  and  $F_{ba}^* = F_{ab}^*$ . In the following, we take  $\mu = 3$  and  $F_{ab}^* = M/|E|$  in all the numerical simulations. Note that the empirical values of  $\mu$  are usually greater than 2 [26, 30]. The selected value of  $\mu$  is to capture the nonlinearity of the relation and it is expected to represent the main qualitative behaviour of the system. The idea behind the above capacities is that the links should nearly behave like free lines when all the movements (here  $M$ ) are uniformly distributed among all the directed edges (here  $|E|$ ). The capacity of a line is of course given by the number of its lanes, but this should somehow depend on the population density (or local density). As mentioned before, in the following we work with a fixed population density  $M/N = 10^3$ . This means that the line capacities are fixed and do not change.

### 3. Results and discussion

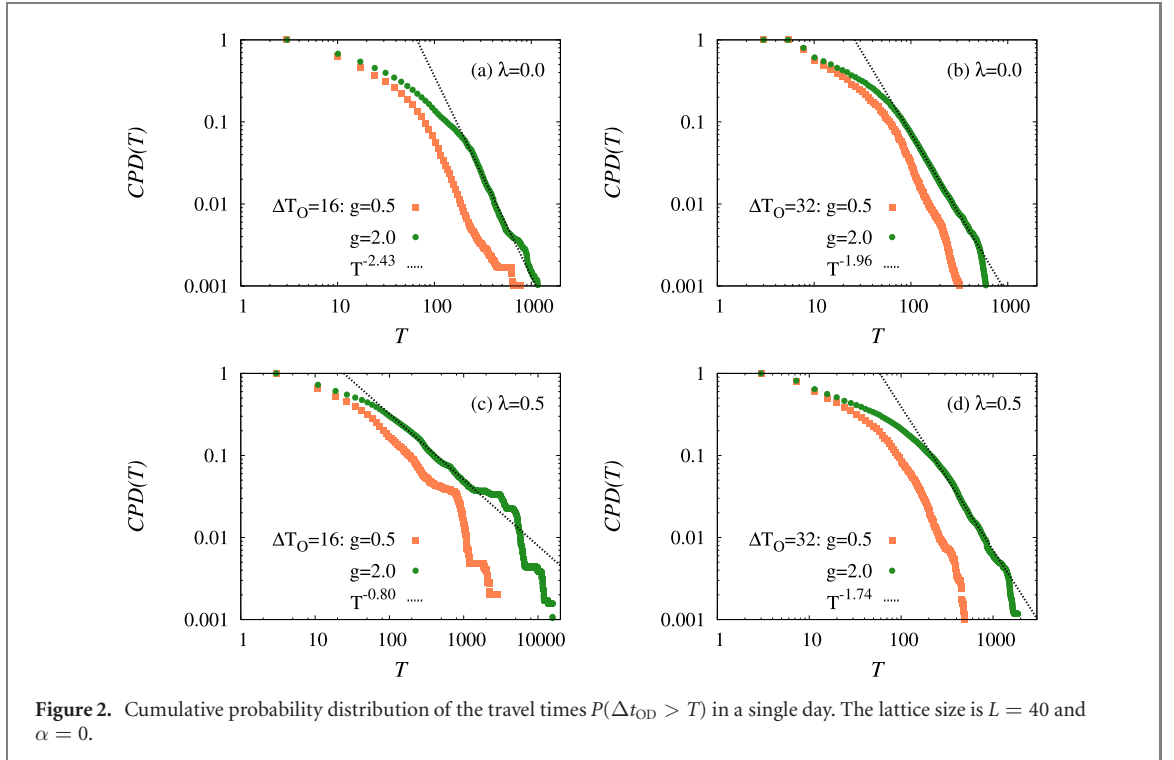
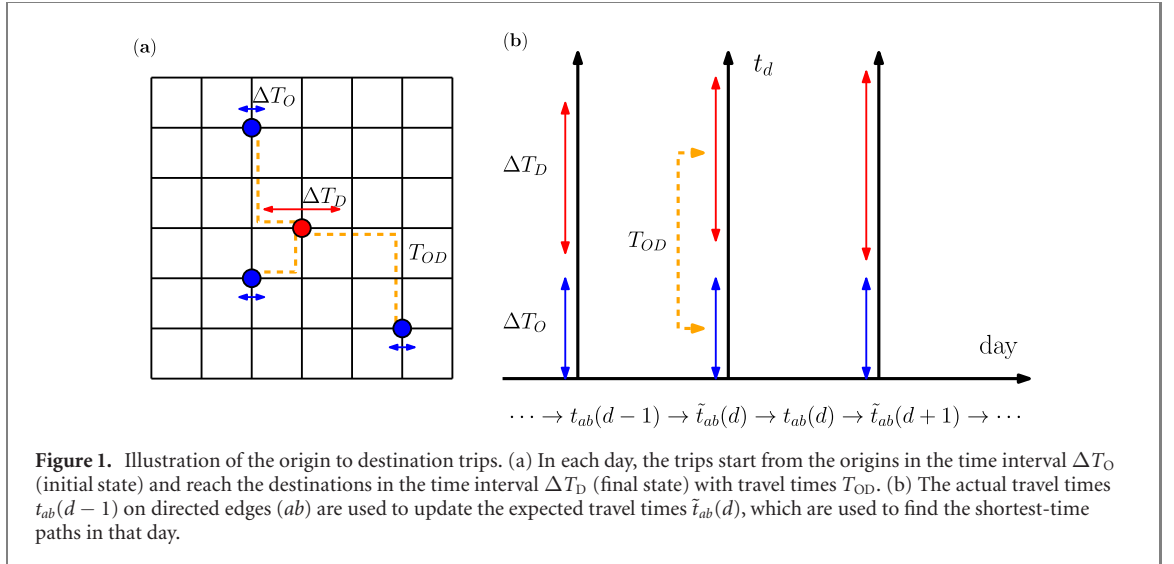
Let us start with the effects of interactions on the cumulative distribution of the travel times  $P(\Delta t_{OD} > T)$  in the absence of any randomness in the movements ( $\alpha = 0$ ). As figure 2 shows, the travel times increase by introducing the interactions either by considering the effects of flows on the trips (with  $g$ ) or by exploiting the travel information from the previous days (with  $\lambda$ ). We know that interactions can reduce predictability in a complex system [31]. Here, it helps to imagine interactions as random scatterings of individuals, which is increasing uncertainty in the arrival times. The stochastic nature of the flows and thus interactions is a simple explanation for the reduction in the predictability of the system. A selfish way of using the travel information without any coordination could even worsen the situation [20]. In the following, we see how the interplay of the above interactions with randomness in the movements affects a measure of predictability of the movement process. In addition, we follow the changes in the efficiency and a measure of entropy production in the system to see how the relation between these quantities depend on the macroscopic state of the system (low-congestion and high-congestion phases) [21].

#### 3.1. Predictability

Predictability of the movements can be studied by characterizing the entropy or mutual information of the relevant quantities [10, 32–34]. For instance, it is interesting to know how much the travel times depend on the geometrical distances and the starting times of the trips. We also study a time series of the travel times in different days to check a measure of predictability in the stationary state of the process.

For a single stochastic variable one may use the (estimated) entropy to quantify the variable uncertainty [35]. A measure of predictability for two stochastic variables  $x, x'$  is provided by the mutual information of the two variables,

$$MI(x, x') = \sum_{x, x'} P(x, x') \log \frac{P(x, x')}{P(x)P(x')}, \quad (4)$$



where  $P(x, x')$ ,  $P(x)$  and  $P(x')$  are the joint and marginalized probability distributions of the variables. Note that large values of mutual information do not necessary mean that accurate prediction is an easy task [36].

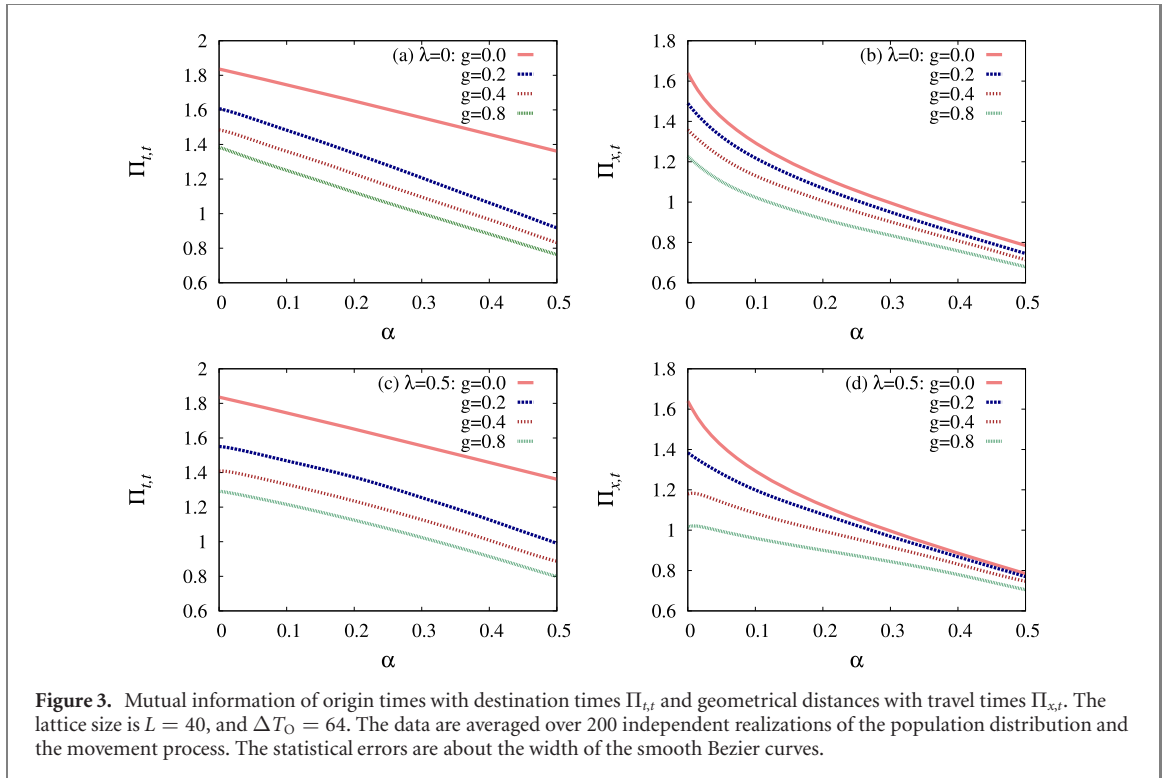
Let us define the mutual information between the destination and origin times in one day (i.e. along the vertical direction in figure 1),

$$\Pi_{t,t} = \text{MI}(t_O, t_D), \quad (5)$$

and the mutual information between the travel time and the geometrical distance,

$$\Pi_{x,t} = \text{MI}(\Delta x_{OD}, \Delta t_{OD}). \quad (6)$$

Figure 3 shows the results of numerical simulations for these quantities in a square lattice of linear size  $L = 40$ . The above measures of predictability diminish with increasing  $g$  and  $\alpha$ , or by using the travel information from the previous days ( $\lambda = 0.5$ ) in the absence of randomness ( $\alpha = 0$ ). The difference between the cases  $\lambda = 0$  and  $\lambda = 0.5$  gets smaller for larger  $\alpha$ , where  $\Pi_{t,t}$  is even a bit larger in the latter case. Note that when  $\lambda = 0$  all edges have the same expected travel times  $\tilde{t}_{ab}(d) = t_{ab}(0) = 1$ , where the shortest path is strongly correlated with the geometrical distance  $\Delta x_{OD}$ . When  $\lambda = 0.5$ , there are some edges with small travel times in the previous day and all drivers are aware of this information which is used to determine the shortest path to their destinations.



Therefore, one expects to observe less correlations here between the geometrical distance and the actual travel times. However, the same global information could make the arrival times at the destinations more dependent on the departure times and results to larger mutual information between the two time variables.

Next, we study the changes in the travel times along the links for different days (i.e. along the horizontal direction in figure 1). Recall that the trips in each day rely on the expected travel times  $\tilde{t}_{ab}(d)$  from the previous days. The actual travel times  $t_{ab}(d)$  in that day could however be different from the estimations we started with. The relative difference of the two quantities thus provides a measure of unpredictability in the movement process,

$$D_d(t, \tilde{t}) = \frac{1}{|E|} \sum_{(ab)} \frac{|t_{ab}(d) - \tilde{t}_{ab}(d)|}{\tilde{t}_{ab}(d)}. \quad (7)$$

Consider a time series of the above quantity in terms of  $d$ . The power spectrum  $S(\omega)$  of such a time series as a function of frequency  $\omega$  is reported in figure 4. Here  $S(\omega) \propto |\hat{D}(\omega)|^2$  where  $\hat{D}(\omega)$  is the discrete Fourier transform of the above time series. The appearance of a dominant peak at high frequencies for larger interactions  $g$  signals a high level of variation at the shortest time scales (two consecutive days). Figure 5 shows the stationary values of  $D_d(t, \tilde{t})$  versus the model parameters  $g$  and  $\lambda$ . Interestingly, the above quantity exhibits a minimum for very small but nonzero values of  $\lambda$  with a discontinuity at  $\lambda = 0$ . This means that some level of information sharing could be helpful even in the absence of any coordination. Moreover, the stronger interactions (larger  $g$ ) result in smaller values and ranges of beneficial  $\lambda$ , which can be used to reduce the above measure of unpredictability. On the other hand, for a fixed  $\lambda$ , the deviations from the expected travel times grow with the strength of interactions as expected.

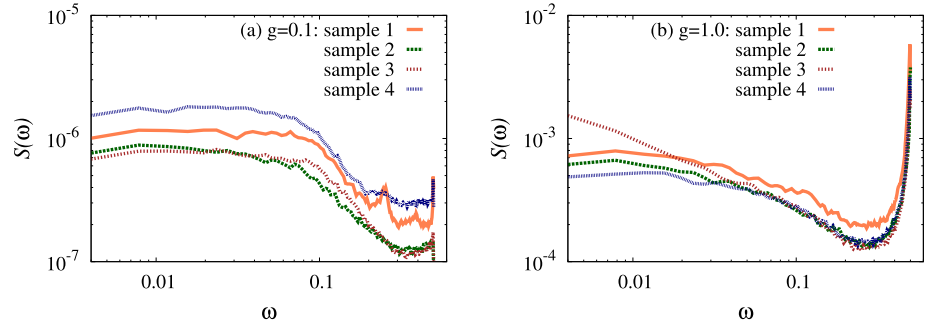
### 3.2. Efficiency and entropy production

Perhaps the most important quantities of the process regarding its efficiency are the travel times and costs. The average travel time in the process is given by

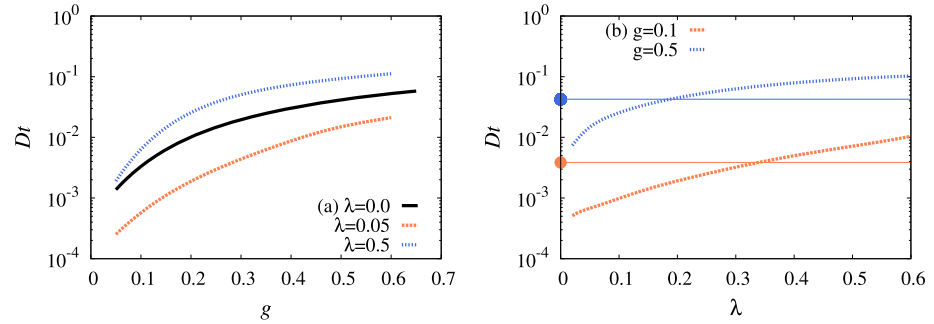
$$\tau_{OD} = \frac{1}{M} \sum_i \Delta t_{OD}(i), \quad (8)$$

where the sum is over all individuals  $i = 1, \dots, M$ . The average number of travels per person in a day is obtained from the sum of all the input flows  $F_{ab}(t)$  for different edges and times,

$$\sigma_{OD} = \frac{1}{M} \sum_t \sum_{(ab)} F_{ab}(t). \quad (9)$$



**Figure 4.** Power spectrum  $S(\omega)$  (square of Fourier transform) of  $D_d(t, \tilde{t})$  time series in a sequence of  $10^5$  days. The lattice size is  $L = 30$ ,  $\Delta T_O = 48$ , and  $\alpha = 0$ . The samples show independent realizations of the population distribution and the movement process. The power was computed by the signal welch method of scientific python (scipy).



**Figure 5.** The average deviation from the expected travel times  $\langle D_d(t, \tilde{t}) \rangle$  in one day. The horizontal lines in panel (b) indicate the values obtained for  $\lambda = 0$ . The lattice size is  $L = 40$ ,  $\Delta T_O = 64$ , and  $\alpha = 0$ . The data are averaged over 200 independent realizations of the population distribution and the movement process in the stationary state. The stationary values are obtained from data samples in an interval of 100 days after a relaxation stage of 200 days. The statistical errors are about the width of the smooth Bezier curves.

This is expected to be proportional to the total energy (e.g. the total amount of fuel) that is consumed in the process. A measure of efficiency then is defined by ratio of the two quantities,

$$\eta_{OD} = \frac{1/\tau_{OD}}{\sigma_{OD}}. \quad (10)$$

This efficiency takes its maximum value for  $g = 0$ ,  $\alpha = 0$ , when both the travel times and the number of edge travels are minimum. For each person  $i$ , we can also define the velocity  $v_{OD}(i) = \Delta x_{OD}(i)/\Delta t_{OD}(i)$ , given the geometrical origin to destination distance  $\Delta x_{OD}(i)$ . Then, the average velocity is

$$v_{OD} = \frac{1}{M} \sum_i v_{OD}(i), \quad (11)$$

which can be regarded as another measure of the process efficiency [37].

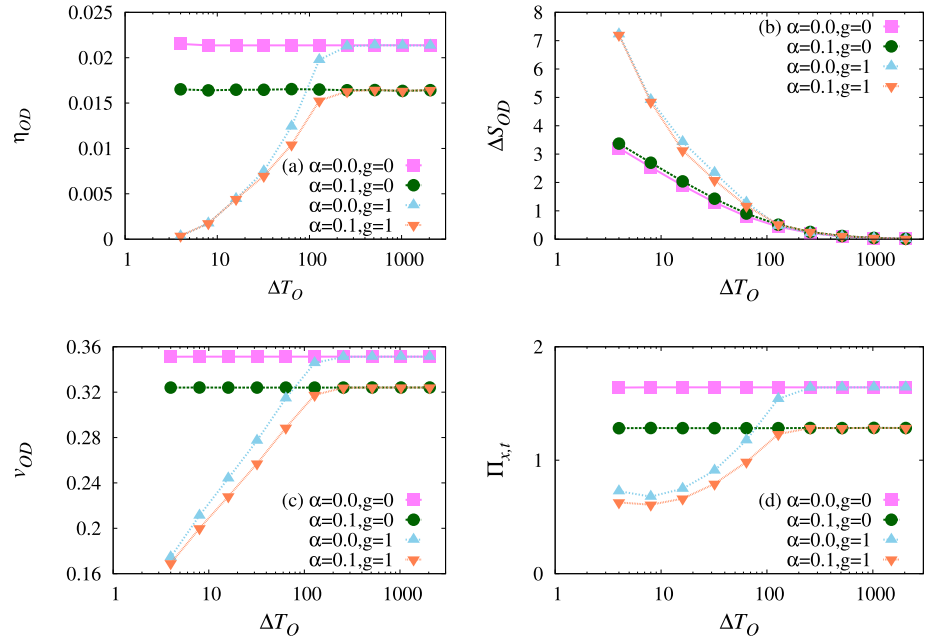
In the following, we are also interested in the amount of disorder that is generated by the movements. A measure of increase in the system disorder or uncertainty (entropy production) is provided by the relative entropy of the origin and destination time intervals,

$$\Delta S_{OD} = \langle \log \Delta T_D \rangle - \langle \log \Delta T_O \rangle = \frac{1}{N} \sum_a \log \Delta T_D(a) - \log \Delta T_O. \quad (12)$$

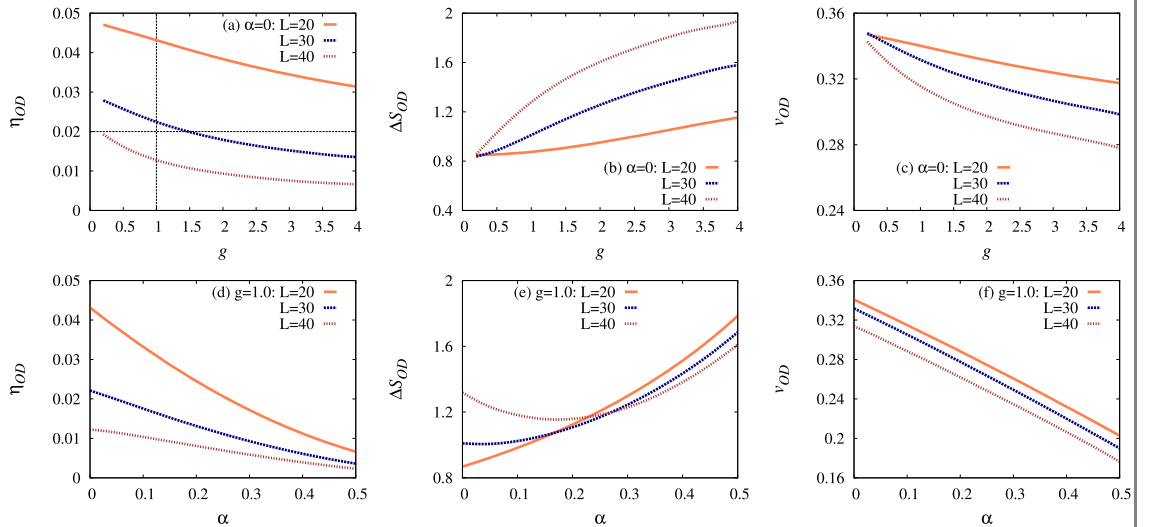
As mentioned before,  $\Delta T_O$  is the same for all sites  $a$ . Other measures of entropy production (or irreversibility) of the process can be defined for example by the relative entropy of the forward (origin to destination) and backward (destination to origin) flows [21]. In this study, we focus on the above measure  $\Delta S_{OD}$ , which only needs the destination time intervals at the end of the forward process, to compare with the origin time intervals.

Figure 6 shows how the above quantities change with the origin time interval  $\Delta T_O$  when  $\lambda = 0$ . We see that by increasing  $\Delta T_O$  (slower process) the efficiency increases because the interaction parameter  $g$  becomes effectively irrelevant for very large time intervals. In this limit, the relative entropy  $\Delta S_{OD}$  is nearly zero but it grows continuously as  $\Delta T_O$  decreases, separating two phases of high and low efficiency (or velocity).





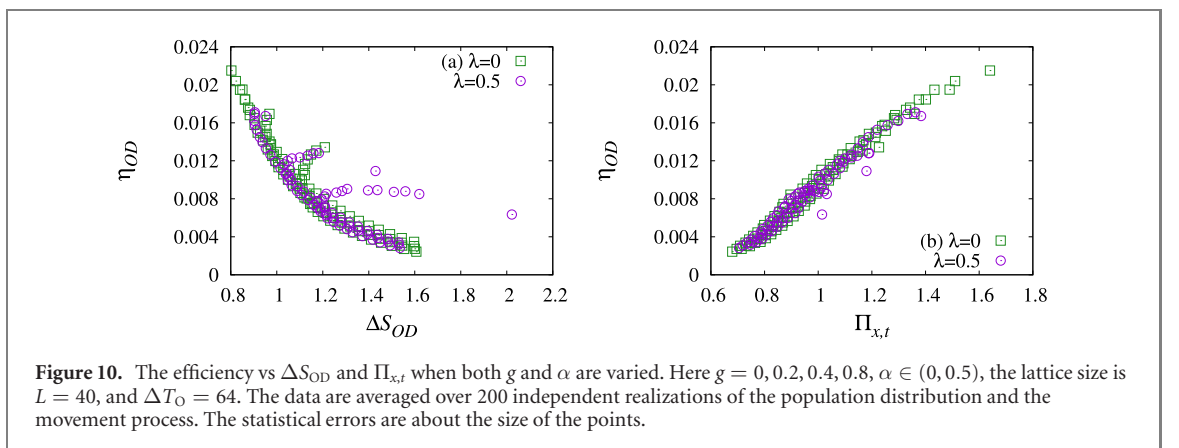
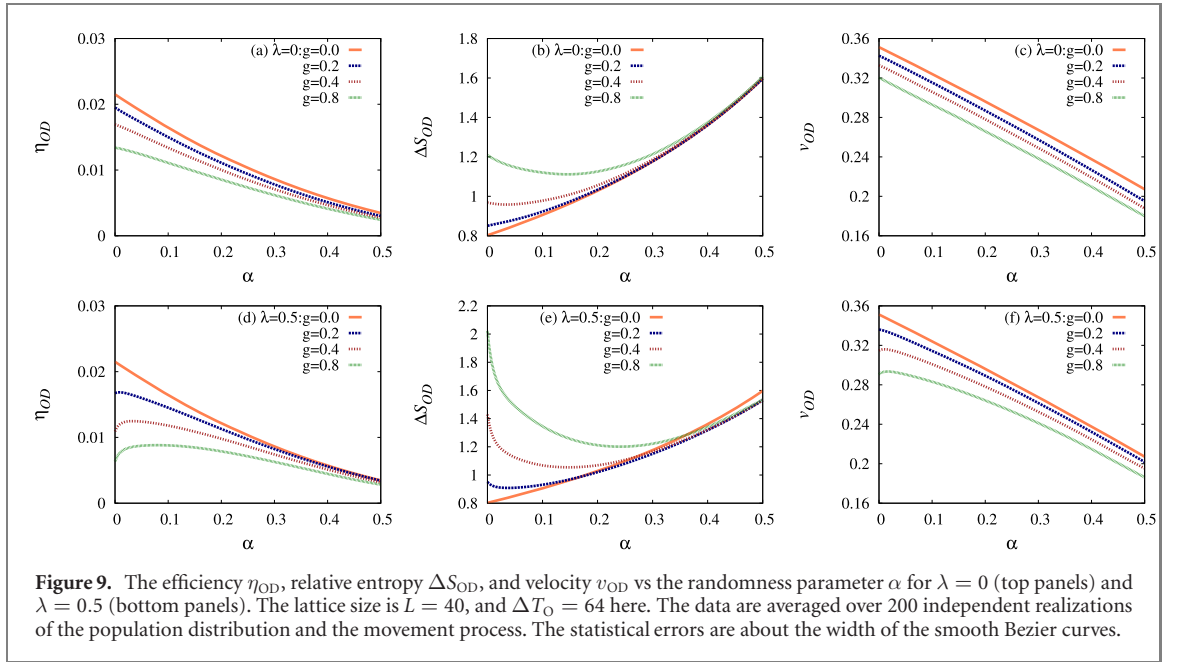
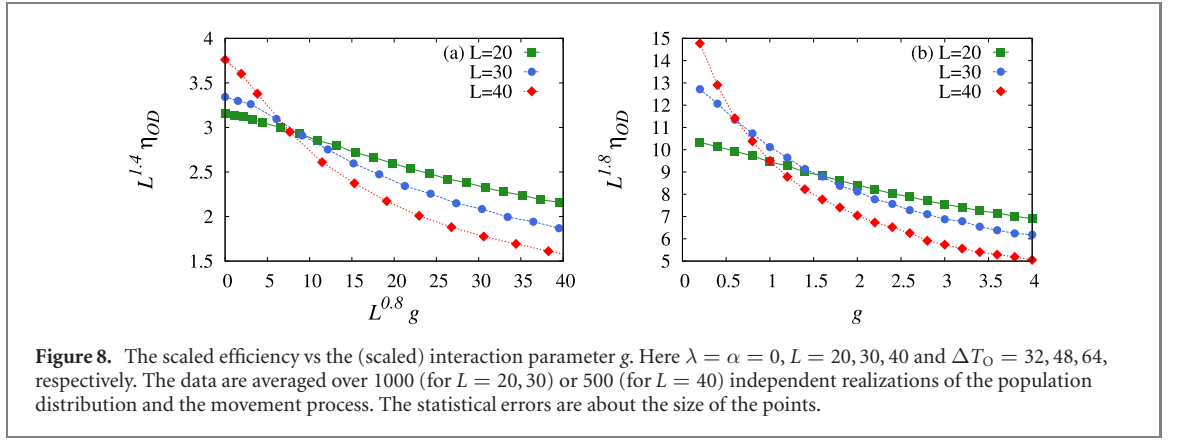
**Figure 6.** The main quantities vs the origin time interval  $\Delta T_O$  for  $\lambda = 0$ . The lattice size is  $L = 40$ . The data are averaged over 200 independent realizations of the population distribution and the movement process. The statistical errors are about the size of the points.



**Figure 7.** Lattice size effects on the main quantities for  $\lambda = 0$ . The selected values of  $\Delta T_O$  and  $L$  are (32, 20), (48, 30), and (64, 40) with fixed ratio  $L/\Delta T_O$ . The population density is fixed to  $M/N = 10^3$ . The data are averaged over 1000 (for  $L = 20, 30$ ) or 500 (for  $L = 40$ ) independent realizations of the population distribution and the movement process. The statistical errors are about the width of the smooth Bezier curves.

In figure 7, we report the lattice size effects when the parameters  $g$  and  $\alpha$  are varied for  $\lambda = 0$ . We keep the population density  $M/N = 1000$  and the links' capacity  $F_{ab}^* = M/|E|$  for different linear sizes  $L = 20, 30, 40$ . It is observed that the efficiency and velocity go to zero very rapidly with  $L$  even when we increase  $\Delta T_O$  to take the ratio  $L/\Delta T_O$  fixed. As figure 8 shows, the efficiency scales roughly as  $L^{-1.8}$  for  $g$  values around 1. This means that, under the above conditions, the city size is limited by the efficiency or velocity of the movement process. A similar phenomenon is observed in [38], where infrastructure costs put an upper bound on the size of a city (in terms of social interactions). The key point is that here the population is increasing with the size and we have a strongly heterogeneous population distribution concentrated around the center of the lattice. More importantly, here the link capacities  $F_{ab}^*$  are the same in whole the lattice and do not change with the size of the system. These assumptions explain why the efficiency approaches zero by increasing  $L$  and why in practice we have to adjust the capacities according to the population size and distribution. The question then





is how we should set the local capacities such that the efficiency approaches a desirable limit as the population grows.

Another interesting observation in figure 7 is that  $\Delta S_{OD}$  initially decreases with the randomness parameter  $\alpha$  and displays a minimum that is more pronounced for larger lattice sizes. This behaviour is better exhibited in figure 9, where we report the results for a fixed  $L = 40$  with different values of  $g$  and  $\lambda$ . In fact, the relative entropy increases with  $\alpha$  for small values of the interaction parameter. But, for  $g$  larger than a critical value, which depends on  $L$  and the other model parameters ( $\Delta T_0$  and  $\lambda$ ), a bit of randomness in the movements

can reduce the relative entropy  $\Delta S_{OD}$ . That is, random deviations from the shortest paths in this regime allow to have more order in the destination time intervals. When  $\lambda = 0.5$  this behaviour is observed for a smaller  $g$ , as using the travel information from the previous days results in a larger effective interaction between the drivers. Besides the minimum displayed by  $\Delta S_{OD}$ , in figure 9 we also observe a maximum in the efficiency and velocity for  $\lambda = 0.5$ . It means that in this case the randomness is also reducing the total travel times by avoiding the congested links of the network. Note that the minimum of entropy production and the maximum of efficiency occur for different values of  $\alpha$ . So, definition of an optimal  $\alpha$  depends on the quantity of interest. Recall that the parameter  $\alpha$  represents the randomness in the local movements of the agents. In practice, such a randomness can be effectively modelled by random local reroutings or by some level of noise in the shared travel information.

In general, the efficiency is positively correlated with the probability of making a right decision which is expected to increase by reducing the travel uncertainties. This relation should be exhibited by reasonable measures of the two quantities (efficiency and uncertainty), see figure 10. As the figure shows, there is however a region where both the efficiency and entropy production decrease with increasing  $\alpha$ . For  $\lambda = 0$ , this happens when randomness in the movements reduces the relative entropy  $\Delta S_{OD}$ . For  $\lambda = 0.5$ , this region shrinks to a smaller interval, that is from the maximum of efficiency to the minimum of  $\Delta S_{OD}$ . In other words, a negative correlation between the efficiency and entropy production is better displayed in the more realistic case when the shortest paths are chosen according to the available travel information.

## 4. Conclusion

In summary, numerical simulations of the movement process shows that in general interactions  $(g, \lambda)$  increase the travel time, reduce the efficiency, and increase the entropy production. A bit of information sharing by exploiting the travel times from the previous days, could reduce the deviation of the actual travel times from the expected ones, so increasing the process predictability. Moreover, some level of randomness in the movements is beneficial for the efficiency (or inverse of the travel time) of selfish drivers in the congested phase. We also observe a qualitative change in behaviour of the entropy production  $\Delta S_{OD}$  with the randomness parameter  $\alpha$  as the parameter  $g$  increases. In fact, for sufficiently large  $g$ , one can reduce the uncertainty in the destination time intervals by introducing randomness in the movements. In other words, the response of  $\Delta S_{OD}$  to the randomness  $\alpha$  can be used to discriminate the low- and high-congestion phases of the flows in the system.

It would be interesting to see how much these findings are sensitive to details of the movements and definition of the efficiency, predictability, and the entropy production of the process. In this study, for instance, we assumed that the departure times are uniformly distributed in  $\Delta T_O$  which can be considered as a first approximation to a unimodal distribution, e.g. a Gaussian one concentrated on a rush hour. However, it is more realistic to assume that distribution of the departure times is not uniform and these variables are correlated with the expected travel times.

The line capacities  $F_{ab}^*$  are another set of important parameters in the problem. Here, we worked with a homogeneous distribution of the capacities that are proportional to the population density  $M/N$ . In this case, the efficiency goes to zero by increasing the system size for a fixed population density. A future study can look for an optimal set of (constrained) line capacities to compensate the reduction in the efficiency as the population grows.

In this paper, we considered selfish drivers which follow the shortest-time paths and studied small deviations from this routing strategy. The other extreme case is the movement of a large number of interacting random walkers where the process is mainly governed by randomness. It would be interesting to see how the entropy production and its relation with efficiency behave in this regime.

## Acknowledgments

This work was performed using the ALICE compute resources provided by Leiden University.

## Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

## ORCID iDs

Abolfazl Ramezanzpour  <https://orcid.org/0000-0002-3650-7722>

## References

- [1] Cresswell T 2006 *On the Move (Mobility in the Modern Western World)* (London: Taylor and Francis)
- [2] Ravenstein E G 1885 *J. Stat. Soc. Lond.* **48** 167
- [3] Stouffer S A 1940 *Am. Socio. Rev.* **5** 845
- [4] Zipf G K 1946 *Am. Socio. Rev.* **11** 677
- [5] Adams J S 2001 *Urban Geogr.* **22** 530
- [6] Ratti C, Frenchman D, Pulselli R M and Williams S 2006 Environment and planning B: urban analytics and city Science **33** 727–48
- [7] Barbosa-Filho H et al 2018 *Phys. Rep.* **734** 1–74
- [8] González M C, Hidalgo C A and Barabási A-L 2008 *Nature* **453** 779
- [9] Song C, Koren T, Wang P and Barabási A-L 2010 *Nat. Phys.* **6** 818
- [10] Song C, Qu Z, Blumm N and Barabási A-L 2010 *Science* **327** 1018
- [11] Li R, Dong L, Zhang J, Wang X, Wang W-X, Di Z and Stanley H E 2017 *Nat. Commun.* **8** 1841
- [12] Alessandretti L, Aslak U and Lehmann S 2020 *Nature* **587** 402
- [13] Wikipedia Contributors 2021 Modal share-Wikipedia, the free encyclopedia [https://en.wikipedia.org/w/index.php?title=Modal\\_share&oldid=1011200149](https://en.wikipedia.org/w/index.php?title=Modal_share&oldid=1011200149) (accessed 10 March 2021).
- [14] Schrank D, Eisele B and Lomax T 2019 Texas A & M Transportation Institute with Cooperation from INRIX The 2019 Urban Mobility Report <https://mobility.tamu.edu/umr/report/>
- [15] Pishue B 2020 2020 global traffic scorecard <https://inrix.com/scorecard>
- [16] Transport statistics Great Britain 2019. Summary. Department for Transport.
- [17] European Environment Agency (EEA) 2020 Luxembourg: Publications Office of the European Union EEA Report, No. 09/2020 p 162 <https://www.eea.europa.eu/publications/air-quality-in-europe-2020-report>
- [18] Horni A, Nagel K and Axhausen K W 2016 *The Multi-Agent Transport Simulation MATSim* (London: Ubiquity Press)
- [19] Kitamura R, Chen C, Pendyala R M and Narayanan R 2000 *Transportation* **27** 25
- [20] Macfarlane J 2019 *IEEE Spectr.* **56** 22
- [21] Biazzo I and Ramezanpour A 2020 *Sci. Rep.* **10** 4334
- [22] Li R, Wang R and Tian X 2020 *J. Phys.: Conf. Ser.* **1550** 032059
- [23] Yan X-Y, Zhao C, Fan Y, Di Z and Wang W-X 2014 *J. R. Soc. Interface* **11** 20140834
- [24] Hardin G 2009 *J. Nat. Resour. Pol. Res.* **1** 243–53
- [25] Roughgarden T 2005 *Selfish Routing and the Price of Anarchy* (Cambridge, MA: MIT Press)
- [26] Youn H, Gastner M T and Jeong H 2008 *Phys. Rev. Lett.* **101** 128701
- [27] U. S. B. of Public Roads 1964 *Traffic Assignment Manual for Application with a Large, High Speed Computer* vol 2 (Washington, DC: US Department of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division, U.S. Government Printing Office)
- [28] Branston D 1976 *Transport. Res.* **10** 223
- [29] Huntsinger L F and Roupail N M 2011 *Transport. Res. Rec.* **2255** 117
- [30] de Dios Ortúzar J and Willumsen L 2011 *Modelling Transport* (New York: Wiley)
- [31] Salganik M J, Dodds P S and Watts D J 2006 *Science* **311** 854
- [32] Gallotti R, Bazzani A, Degli Esposti M and Rambaldi S 2013 *J. Stat. Mech.* **P10022**
- [33] Lu X, Wetter E, Bharti N, Tatem A J and Bengtsson L 2013 *Sci. Rep.* **3** 2923
- [34] Wang H, Zeng S, Li Y and Jin D 2020 *IEEE Trans. Mobile Comput.* **20** 2457–72
- [35] Fano R M 1949 *The Transmission of Information* (Massachusetts Institute of Technology Research Laboratory of Electronics) Technical Report No. 65
- [36] Navet N and Chen S-H 2008 *Natural Computing in Computational Finance* (Berlin: Springer) pp 197–210
- [37] Biazzo I, Monechi B and Loreto V 2019 *R. Soc. Open Sci.* **6** 190979
- [38] Bettencourt L M A 2013 *Science* **340** 1438